U.S. DEPARTMENT OF COMMERCE
National Technical Information Service

PB-254 658

Quality Assurance Handbook for Air Pollution Measurement Systems Volume I. Principles

Environmental Monitoring & Support Lab.

Prepared For
PEDCo-Environmental Specialists, Inc.

March 1976

KEEP UP TO DATE

Between the time you ordered this report—which is only one of the hundreds of thousands in the NTIS information collection available to you—and the time you are reading this message, several *new* reports relevant to your interests probably have entered the collection.

Subscribe to the **Weekly Government Abstracts** series that will bring you summaries of new reports *as soon as they are received by NTIS* from the originators of the research. The WGA's are an NTIS weekly newsletter service covering the most recent research findings in 25 areas of industrial, technological, and sociological interest—invaluable information for executives and professionals who must keep up to date.

The executive and professional information service provided by NTIS in the **Weekly Government Abstracts** newsletters will give you thorough and comprehensive coverage of government-conducted or sponsored research activities. And you'll get this important information within two weeks of the time it's released by originating agencies.

WGA newsletters are computer produced and electronically photocomposed to slash the time gap between the release of a report and its availability. You can learn about technical innovations immediately—and use them in the most meaningful and productive ways possible for your organization. Please request NTIS-PR-205/PCW for more information.

The weekly newsletter series will keep you current. But *learn what you have missed in the past* by ordering a computer **NTISearch** of all the research reports in your area of interest, dating as far back as 1964, if you wish. Please request NTIS-PR-186/PCN for more information.

WRITE: Managing Editor
5285 Port Royal Road
Springfield, VA 22161

Keep Up To Date With SRIM

SRIM (Selected Research in Microfiche) provides you with regular, automatic distribution of the complete texts of NTIS research reports *only* in the subject areas you select. SRIM covers almost all Government research reports by subject area and/or the originating Federal or local government agency. You may subscribe by any category or subcategory of our WGA **(Weekly Government Abstracts)** or **Government Reports Announcements and Index** categories, or to the reports issued by a particular agency such as the Department of Defense, Federal Energy Administration, or Environmental Protection Agency. Other options that will give you greater selectivity are available on request.

The cost of SRIM service is only 45¢ domestic (60¢ foreign) for each complete microfiched report. Your SRIM service begins as soon as your order is received and processed and you will receive biweekly shipments thereafter. If you wish, your service will be backdated to furnish you microfiche of reports issued earlier.

Because of contractual arrangements with several Special Technology Groups, not all NTIS reports are distributed in the SRIM program. You will receive a notice in your microfiche shipments identifying the exceptionally priced reports not available through SRIM.

A deposit account with NTIS is required before this service can be initiated. If you have specific questions concerning this service, please call (703) 451-1558, or write NTIS, attention SRIM Product Manager.

This information product distributed by

NTIS U.S. DEPARTMENT OF COMMERCE
National Technical Information Service
5285 Port Royal Road
Springfield, Virginia 22161

183096

EPA-600/9-76-005
January 1976

PB 254 658

QUALITY ASSURANCE HANDBOOK FOR AIR POLLUTION MEASUREMENT SYSTEMS

Volume I — Principles

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U. S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

U.S. ENVIRONMENTAL PROTECTION AGENCY
Office of Research and Development
Environmental Monitoring and Support Laboratory
Research Triangle Park, North Carolina 27711

379

ACKNOWLEDGMENTS

This volume of the Quality Assurance Handbook has been prepared by the quality assurance staff of the Environmental Monitoring ans Support Laboratory, Research Triangle Park, North Carolina in cooperation with PEDCo-Environmental Specialists, Inc., of Cincinnati, Ohio.

DISCLAIMER

Mention of trade names or commercial products does not constitute EPA endorsement or recommendation for use.

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA-600/9-76-005	2.	3. RECIPIENT'S ACCESSION NO.

4. TITLE AND SUBTITLE QUALITY ASSURANCE HANDBOOK FOR AIR POLLUTION MEASUREMENT SYSTEMS, VOLUME I - PRINCIPLES.	5. REPORT DATE March 1976
	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Darryl J. von Lehmden, Environmental Protection Agency Carl Nelson, PEDCO, Environmental Specialists, Inc.	8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS U.S. Environmental Protection Agency, Environmental Monitoring and Support Laboratory, Research Triangle Park, N. C. and PEDCO Environmental Specialists, Inc., Cincinnati, Ohio.	10. PROGRAM ELEMENT NO. 1HA327
	11. CONTRACT/GRANT NO. 68-02-1343. Task No. 5
12. SPONSORING AGENCY NAME AND ADDRESS Environmental Monitoring and Support Laboratory U.S. Environmental Protection Agency Office of Research and Development Research Triangle Park, North Carolina 27711	13. TYPE OF REPORT AND PERIOD COVERED Handbook
	14. SPONSORING AGENCY CODE EPA-ORD

15. SUPPLEMENTARY NOTES The Handbook will be reproduced in the current format (3-ring binder) and maintained up-to-date by a document control system operated by EMSL/RTP. Distribution will be to personnel in EPA, and its contractors, state and selected air pollution control agencies in foreign countries.

16. ABSTRACT

This Handbook includes guiding principles and recommended procedures for achievin quality assurance in air pollution measurement systems. Vol. I contains brief discussions of elements of quality assurance. Expanded discussions of technical points, including sample calculations, are included in the appendices. The discussion of each element, therefore, is structured to be brief and at the same time, to highlight its most important features. The elements of quality assurance discussed in Vol. I include document control, quality assurance policy and objectives, organization, quality planning, training, pretest preparation, preventive maintenance, sample collection, sample analysis data reporting, procurement quality control, calibration, corrective action, quality cost interlaboratory and intralaboratory testing, audit procedures, data validation, statistical analysis of data, configuration control, reliability, quality reports to management, quality assurance manual and quality assurance plans fo projects and programs. Organizations developing and implementing their own quality assurance plans will find Vol. I useful for general guidance.

PRICES SUBJECT TO CHANGE

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Quality assurance	Interlaboratory testing	13B
Quality Control	Configuration control	14B
Preventive maintenance	Document control	
Calibrating	Performance & system audits	
	Procurement quality control	
	Quality assurance plan	
iia	Quality assurance glossary	

18. DISTRIBUTION STATEMENT Release Unlimited	19. SECURITY CLASS *(This Report)* Unclassified	21. NO. OF PAGES 379
	20. SECURITY CLASS *(This page)* Unclassified	22. PRICE $10.75 - 2.25

EPA Form 2220-1 (9-73)

VOLUME I.

TABLE OF CONTENTS

Section		Pages	Revision	Date
1.1	PURPOSE OF THE QUALITY ASSURANCE HANDBOOK	2	0	5-1-75
1.2	OVERVIEW OF THE QUALITY ASSURANCE HANDBOOK	4	0	5-1-75
1.3	DEFINITION OF QUALITY ASSURANCE	2	0	5-1-75
1.4	ELEMENTS OF QUALITY ASSURANCE	3	0	5-1-75
1.4.1	Document Control and Revisions	5	0	5-1-75
1.4.2	Quality Assurance Policy and Objectives	5	0	5-1-75
1.4.3	Organization	7	0	5-1-75
1.4.4	Quality Planning	6	0	5-1-75
1.4.5	Training	8	0	5-1-75
1.4.6	Pre-Test Preparation	6	0	5-1-75
1.4.7	Preventive Maintenance	8	0	5-1-75
1.4.8	Sample Collection	2	0	5-1-75
1.4.9	Sample Analysis	5	0	5-1-75
1.4.10	Data Reporting	10	0	5-1-75
1.4.11	Procurement Quality Control	15	0	5-1-75
1.4.12	Calibration	18	0	5-1-75

TABLE OF CONTENTS (continued)

Section		Pages	Revision	Date
1.4.13	Corrective Action	9	0	5-1-75
1.4.14	Quality Costs	7	0	5-1-75
1.4.15	Interlaboratory and Intralaboratory Testing	11	0	5-1-75
1.4.16	Audit Procedures	4	0	5-1-75
1.4.17	Data Validation	13	0	5-1-75
1.4.18	Statistical Analysis of Data	3	0	5-1-75
1.4.19	Configuration Control	5	0	5-1-75
1.4.20	Reliability	7	0	5-1-75
1.4.21	Quality Reports to Management	5	0	5-1-75
1.4.22	Quality Assurance Manual	2	0	5-1-75
1.4.23	Quality Assurance Plans for Projects and Programs	7	0	5-1-75

APPENDICES

		Pages	Revision	Date
A	GLOSSARY OF TERMS	28	0	5-1-75
B	NOMENCLATURE	7	0	5-1-75
C	COMPUTATIONAL EXAMPLES OF DESCRIPTIVE STATISTICS Data presentation, frequency distribution, measures of central tendency (arithmetic mean, median, geometric mean) measures of dispersion (range, variance and standards deviation, geometric standard deviation, use of range to estimate standard deviation, coefficient of variation, percent differences), computational	25	0	5-1-75

APPENDICES (continued).

Section	Pages	Revision	Date
procedures using coded data, instructions for calculations			
D PROBABILITY DISTRIBUTIONS	16	0	5-1-75
Normal distribution, lognormal distribution, distribution of sample means, use of probability graph paper			
E ESTIMATION PROCEDURES	15	0	5-1-75
Estimation (point, estimate, confidence interval, t distribution), determination of sample size			
F OUTLIERS	10	0	5-1-75
Procedures for identifying outliers			
G TREATMENT OF AUDIT DATA	9	0	5-1-75
Audit data, analysis of audit data (criterion 1, criterion 2 - sign test, paired + - test - criterion 3)			
H CONTROL CHARTS	29	0	5-1-75
Description and theory, applications and limitations, construction of control charts (precision, accuracy, percent of defective measurements, individual results, moving averages and ranges, other charts), interpretation of control charts, steps in developing and using control charts			

APPENDICES (continued).

Section		Pages	Revision	Date
I	STATISTICAL SAMPLING Concepts, how does one select a random sample, stratified random sample? (random sample, stratified sample) acceptance sampling by attributes, acceptance sampling by variables	11	0	5-1-75
J	CALIBRATION Concepts, multipoint calibration, zero-span calibration, suggested calibration procedure	24	0	5-1-75
K	INTERLABORATORY TESTING Concepts, interlaboratory test, data analysis (analysis of interlaboratory tests - summary analysis, analysis for a particular lab, feedback of information on methods - analysis of collaborative test data)	10	0	5-1-75
L	RELIABILITY AND MAINTAINABILITY Life testing	9	0	5-1-75

1.1 PURPOSE OF THE QUALITY ASSURANCE HANDBOOK

The purpose of this Quality Assurance Handbook for Air Pollution Measurement Systems is to provide guidelines and procedures for establishing a new or improving an existing quality assurance program. It is intended to serve as a resource document for the design of quality assurance programs and to provide detailed operational procedures for certain measurement processes that can be used directly in implementing the quality assurance program. This Handbook should be particularly beneficial to operators, project officers, and program managers responsible for implementing, designing and coordinating air pollution monitoring projects.

The Handbook is a compilation of quality assurance principles, practices, guidelines, and procedures that are applicable to air pollution measurement systems.

What is presented in the Handbook is an "ultimate" or "ideal" quality assurance program for air pollution measurement systems. All specific measurement systems will not be able to follow all the principles and guidelines contained in the Handbook. A quality assurance program for air pollution measurement systems should cover a number of areas or elements. These elements are shown in Figure 1.4.1 in a "Quality Assurance Wheel." The wheel arrangement illustrates the need for a quality assurance system that will address all elements and at the same time will allow program managers the flexibility to emphasize those elements that are most applicable to their particular program.

1.2 OVERVIEW OF THE QUALITY ASSURANCE HANDBOOK

This Handbook includes guiding principles and recommended procedures for achieving quality assurance in air pollution measurement systems. It provides general guidelines applicable to most measurement systems as well as specific guidelines applicable to particular measurement processes.

Volume I contains brief discussions of the elements of quality assurance. Expanded discussions of technical points, including sample calculations, are included in the Appendices. The discussion of each element, therefore, is structured to be brief and at the same time to highlight its most important features. Organizations developing and implementing their own quality assurance plans will find Volume I useful for general guidance.

Volume II contains quality assurance guidelines on ambient air measurement systems. Regardless of the scope and magnitude of ambient air measurement systems, there are a number of common considerations pertinent to the production of quality data. These common parameters are discussed in Section 2.0 of Volume II and include quality assurance guidelines in the areas of:

1. Sampling site selection considerations - network design, representative sampling consistent with objectives, meteorological and topographical restraints, and sampling schedules.

2. Sampling considerations - environmental controls, probe and manifold design, and support services.
3. Data handling and reporting considerations - data recording systems, data validation, and systematic data management.

The remainder of Volume II, contains pollutant-specific quality assurance guidelines in the following areas:

1. Measurement method description - step-by-step description of sampling, analysis, calibration and other important measurement activities.
2. Method activity matrix - tabulated synopsis of important characteristics or sources of variation in the measurement method including acceptance limits, frequency of checking the characteristics, recommended measurement method, and recommended action if acceptance limits are not met.
3. Operational procedures - step-by-step procedures and operational checks for the pollutant-specific measurement method.
4. Auditing procedures - recommended audits that should be performed during sampling, analysis, data processing, and any other activity that has been identified as a potential source of error.

Initially Volume II will contain three pollutant-specific measurement systems. These are:

1. Reference Method for the Determination of Sulfur Dioxide in the Atmosphere (Pararosaniline Method).
2. Reference Method for the Continuous Measurement of Carbon Monoxide in the Atmosphere (Nondispersive Infrared Spectrometry).
3. Reference Method for the Determination of Suspended Particulates in the Atmosphere (High Volume Method).

As additional methods are added to Volume II, these will be sent to Handbook holders through the document control system used for this Handbook.

As of the original writing of this Handbook, plans are underway for the development of Volume III, Stationary Source Specific Methods. The format for Volume III will be patterned after that of Volume II, Ambient Air Specific Methods.

Regardless of the scope and purpose of the emissions testing planned, there are a number of common considerations pertinent to the production of quality data. These common parameters are discussed in Section 3.0 of Volume III and include quality assurance guidelines in the area of:

1. Planning the test

2. General source testing factors
3. Data evaluation

The remainder of Volume III contains pollutant-specific quality assurance guidelines. The first pollutant-specific measurement systems to be included in Volume III will be:

Method 2 - Determination of Stack Gas Velocity and Volumetric Flow Rate (Type S Pitot Tube).

Method 3 - Gas Analysis for Carbon Dioxide, Oxygen, Excess Air, and Dry Molecular Weight.

Method 4 - Determination of Moisture in Stack Gases.

Method 5 - Determination of Particulate Emissions from Stationary Sources.

Method 6 - Determination of Sulfur Dioxide Emissions from Stationary Sources.

Method 7 - Determination of Nitrogen Oxide Emissions from Stationary Sources.

Method 8 - Determination of Sulfuric Acid Mist and Sulfur Dioxide.

1.3 DEFINITION OF QUALITY ASSURANCE$^{(1-4)}$

Quality assurance and quality control have been defined and interpreted in many ways. The more authoritative usages differentiate between the two terms by stating that quality control is "the system of activities to provide a quality product," whereas quality assurance is "the system of activities to provide assurance that the quality control system is performing adequately." In other words, quality assurance is quality control for quality control.

Quality control may also be understood as "internal quality control;" namely, routine checks included in normal internal procedures; e.g., periodic calibrations, duplicate checks, split samples, and spiked samples. Quality assurance may also be viewed as "external quality control," those activities that are performed on a more occasional basis, usually by a person outside of the normal routine operations; e.g., on-site system surveys, independent performance audits, interlaboratory comparisons, and periodic evaluation of internal quality control data.

In this Handbook, the term quality assurance is used to include the above meanings of both quality assurance and quality control.

While the objective of EPA's air programs is to improve the quality of the air, the objective of quality assurance for air programs is to improve or assure the quality of measured data on pollutant concentrations. Thus the "product" with which quality assurance is concerned is data.

Since air pollution measurements are made by numerous agencies and private organizations at a large number of field stations and laboratories, quality assurance is also concerned with establishing and assessing comparability of data quality among organizations contributing to data bases.

1.3.1 REFERENCES

1. Juran, J.M. Quality Control Handbook, 3rd Ed. McGraw-Hill, 1974. Section 2.

2. ASTM. Designation E-36, "Standard Recommended Practice for Generic Criteria for Use in the Evaluation of Testing and/or Inspection Agencies."

3. ASQC. Standard A3-1971 (ANSI Standard Z1.7-1971), "Glossary of General Terms Used in Quality Control."

4. Feigenbaum, A.V. Total Quality Control, Engineering and Management. McGraw-Hill, 1961.

1.4 ELEMENTS OF QUALITY ASSURANCE

A quality assurance program for air pollution measurement systems should cover a number of areas or elements. These elements are shown in Figure 1.4.1 in a "Quality Assurance Wheel." The wheel arrangement illustrates the need for a quality assurance system that will address all elements and at the same time will allow program managers the flexibility to emphasize those elements that are most applicable to their particular program. Quality assurance elements are grouped on the wheel according to the organization level to which responsibility is normally assigned. These organizational levels are the quality assurance coordinator (normally a staff function), supervisor (a line function), and the operator. Together the supervisor and quality assurance coordinator must see that all these elements form a complete and integrated system and are working to achieve the desired program objectives.

The three-digit numbers shown on the wheel show the location in Section 1.4 where a description of the element is provided. Each element is described in three subsections as follows:

1. ABSTRACT - A brief summary that allows the program manager to review the section at a glance.
2. DISCUSSION - Detailed text that expands on items summarized in the ABSTRACT.

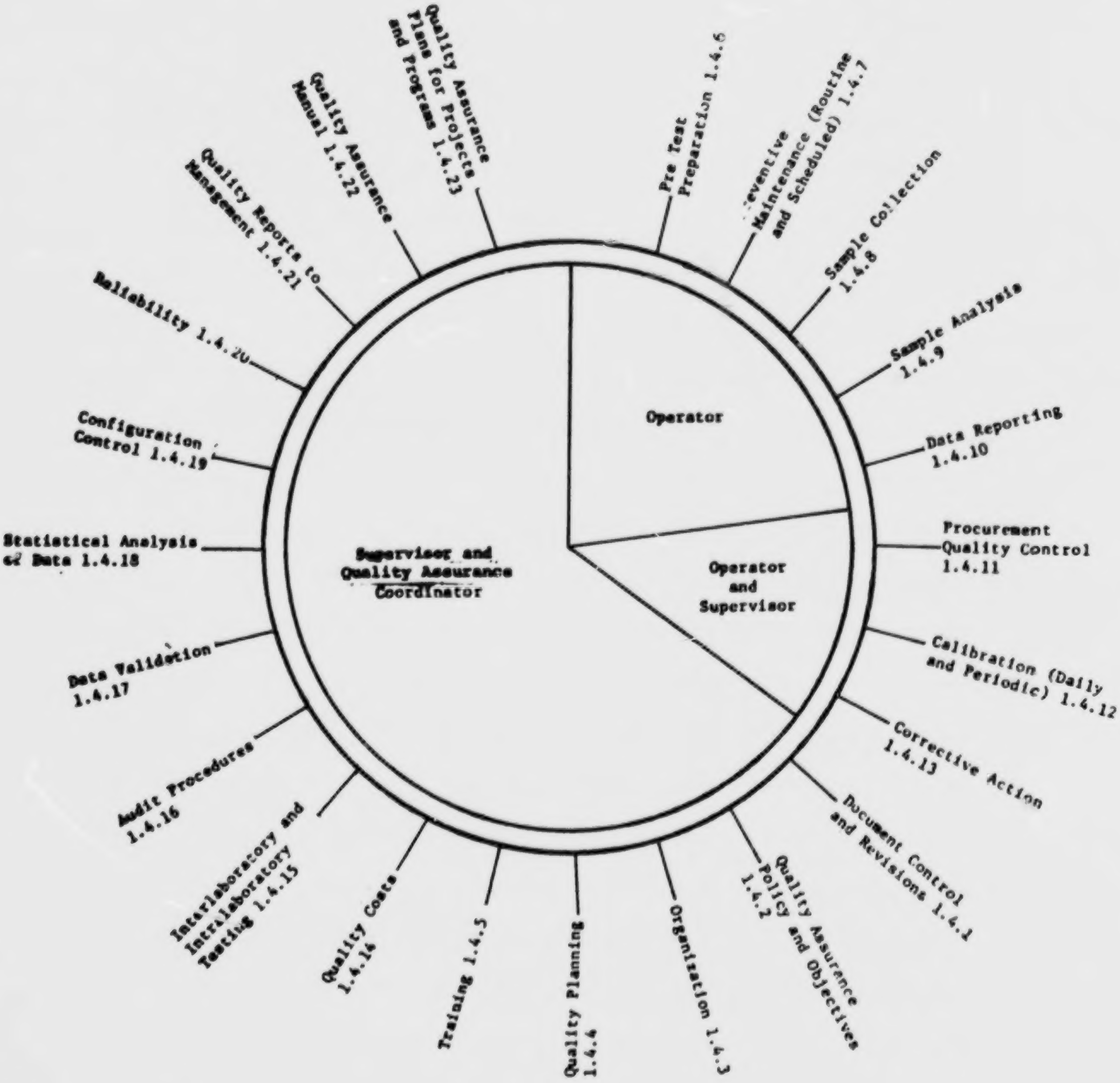

Operator
Operator and Supervisor
Supervisor and Quality Assurance Coordinator
Document Control and Revisions 1.4.1
Quality Assurance Policy and Objectives 1.4.2
Organization 1.4.3
Quality Planning 1.4.4
Training 1.4.5
Pre Test Preparation 1.4.6
Preventive Maintenance (Routine and Scheduled) 1.4.7
Sample Collection 1.4.8
Sample Analysis 1.4.9
Data Reporting 1.4.10
Procurement Quality Control 1.4.11
Calibration (Daily and Periodic) 1.4.12
Corrective Action 1.4.13
Quality Costs 1.4.14
Interlaboratory and Intralaboratory Testing 1.4.15
Audit Procedures 1.4.16
Data Validation 1.4.17
Statistical Analysis of Data 1.4.18
Configuration Control 1.4.19
Reliability 1.4.20
Quality Reports to Management 1.4.21
Quality Assurance Manual 1.4.22
Quality Assurance Plans for Projects and Programs 1.4.23

Figure 1.4.1 Quality assurance elements and responsibilities (The quality assurance wheel)

3. REFERENCES - List of resource documents used in preparation of the discussion. In addition, where applicable, a list of resource documents for recommended reading is shown under BIBLIOGRAPHY.

The DISCUSSION subsection is designed to be relatively brief. In those cases where a topic would require considerable detailed discussion, the reader is referred to the APPENDIX. A case in point is Section 1.4.10 on Data Reporting. In this section, it is recommended that Quality Control Charts be used to monitor performance of critical operations in a measurement system. A brief description of control charts is shown in Section 1.4.10. For more details on types of control charts, step-by-step procedure for preparation, interpretation of control charts, etc., the reader is referred to APPENDIX H - Control Charts.

1.4.1 DOCUMENT CONTROL AND REVISIONS

1.4.1.1 ABSTRACT

A quality assurance program should include a system for documenting operating procedures and revisions in the procedures. The system used for this Handbook is described and is recommended.

1.4.1.2 DISCUSSION

A quality assurance program should include a system for updating formal documentation of operating procedures. The suggested system is the one used in this Handbook and described herein. This system uses a standardized indexing format and provides for convenient replacement of pages that may be changed within the technical procedure descriptions.

The indexing format includes, at the top of each page, the following information:

Section No.
Revision No.
Date (of revision)
Page

A digital numbering system identifies sections within the text. The "Section No." at the top of each page identifies major three-digit or two-digit sections, where applicable. For example, Section 1.4.4 represents "Quality Planning" and

Section 1.4.5 represents "Training." "Revision No." represents the most current version of the section in question, where the first version is represented as "0." "Date" represents the date of the latest revision. "Page No." includes not only the number of the specific page, but also the total number of pages in this section. An example of the page label for the first page of "Quality Planning" in Section 1.4.4 follows:

Section No. 1.4.4

Revision No. 0

Date May 1, 1975

Page 1 of 6

For each three-digit level, the text begins on a new page. This format groups the pages together to allow convenient revision of the section. Each time a new page is added or expanded within a section, the number of the preceding or original page is included on the new page, and a letter is added to it. For example, if Page 4 of 8 were revised and expanded to include an extra paragraph, the overflow would appear on a page designated 4a. The original Page 4 would then be removed from the Handbook and replaced by revised Page 4 and Page 4a. This allows expansion within a section without retyping the section or renumbering all of the pages.

The Table of Contents follows the same structure as the text. It contains a space for "Revision No." and "Pages" within each section heading. When a revision to the text is made, the Table of Contents page would be updated by retyping, or by striking out the old revision number and printing the current revision number. For example, the Table of Contents page detailing Section 1.4 might appear as follows:

		Pages	Revision	Date
1.4.1	Document Control and Revisions	5	0	5/1/75
1.4.2	Quality Assurance Policy and Objectives	5	0	5/1/75
1.4.3	Organization	7	0	5/1/75

A revision to "Organization" would change the Table of Contents to appear as follows:

		Pages	Revision	Date
1.4.1	Document Control and Revisions	5	0	5/1/75
1.4.2	Quality Assurance Policy and Objectives	5	0	5/1/75
1.4.3	Organization	7	1	7/1/75

A Handbook distribution record has been established and will be maintained up to date so that future versions of existing Handbook sections and the addition of new sections may be distributed to Handbook users. In order to enter the user's name and address in the distribution record system, the "Distribution Record Card" in the front of Volume I of this Handbook must be filled out and mailed in the preaddressed envelope provided with this volume of the Handbook. (Note: A separate card must be filled out for each volume of the Handbook). Any future change in name and/or address should be sent to the following:

U.S. Environmental Protection Agency
Environmental Monitoring and Support Laboratory
Quality Assurance Branch
Research Triangle Park, N.C. 27711

Attn: Distribution Record System

Changes may be made by the issuance of (1) an entirely new document, (2) replacement pages or, (3) an errata sheet or by pen and ink posting on the original document. The recipient of these changes should remove and destroy all revised pages, sections, etc. from his copy and/or make pen and ink corrections as noted.

The document control system described herein applies to this Handbook and it can be used, with minor revisions, to maintain control of quality assurance procedures developed by users of this Handbook and quality assurance coordinators. The most important elements of the quality assurance program to which document control should be applied include:

1. Sampling procedures.
2. Calibration procedures.
3. Analytical procedures.

4. Data collection and reporting procedures.
5. Auditing procedures.
6. Sample shipping and storage procedures.
7. Computational and data validation procedures.
8. Quality assurance manuals.
9. Quality assurance plans.

1.4.1.3 REFERENCE

1. Industrial Hygiene Service Laboratory Quality Control Manual. DHEW, PHS, National Institute of Occupational Safety and Health, Cincinnati, Ohio. Technical Report No. 78, 1974. Section VII.

1.4.2 QUALITY ASSURANCE POLICY AND OBJECTIVES

1.4.2.1 ABSTRACT

1. Each organization should have a written quality assurance policy that should be made known to all organization personnel.

2. The objectives of quality assurance are to produce data that meet users' requirements measured in terms of completeness, precision, accuracy, representativeness and comparability and at the same time reduce quality costs.

1.4.2.2 DISCUSSION

Quality assurance policy - Each organization should have a written quality assurance policy. This policy should be distributed so that all organization personnel know the policy and scope of coverage.

Quality assurance objectives[1,2] - To administer a quality assurance program, the objectives of the program must be defined, documented, and issued to all activities that affect the quality of the data. Such written objectives are needed because they:

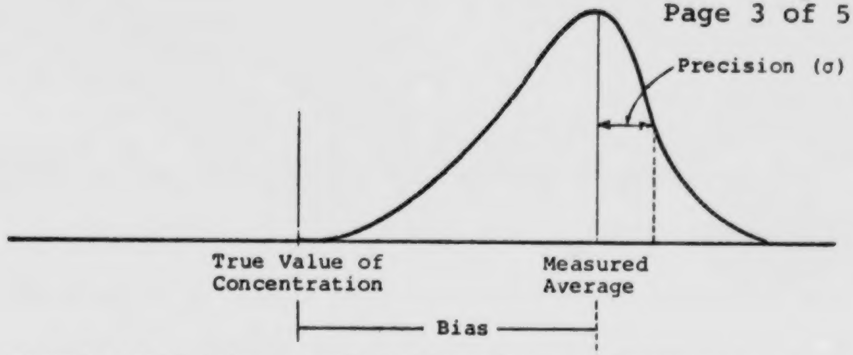
1. Unify the thinking of those concerned with quality assurance.
2. Stimulate effective action.
3. Are a necessary prerequisite to an integrated, planned course of action.

4. Permit comparison of completed performances against stated objectives.

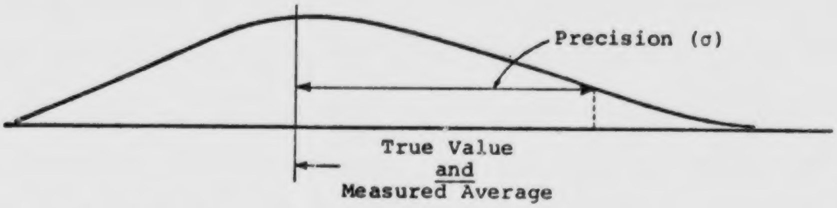
The objectives of a quality assurance system as described in this Handbook are to produce data that are complete, precise, accurate, representative, and comparable.

Data can be considered to be complete if a prescribed percentage of the total possible measurements is present. Precision and accuracy represent measures of the data quality. Data must be representative of the condition being measured. Ambient air sampling at midnight is not representative of carbon monoxide levels during rush hour traffic. Stationary source emission measurements are not representative if measured at reduced load production conditions when normal operation is at full load. Data available from numerous agencies and private organizations should be in consistent units and should be corrected to the same standard conditions of temperature and pressure to allow comparability of data among groups.

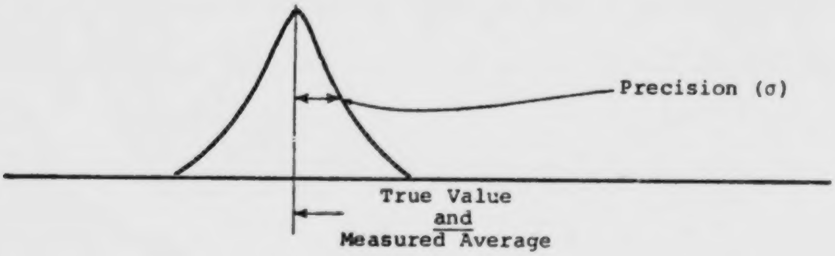
Figure 1.4.2.1 shows three examples of data quality with varying degrees of precision and accuracy. These examples hypothesize a true value that would result if an accurate measurement procedure were available and an infinitely large number of measurements could be made under specified conditions. If the average value coincides with the true value (reference standard), then the measurements are accurate. If the measurement values also are closely clustered about the true value, the measurements are both


Precision (σ)
True Value of Concentration
Measured Average
Bias

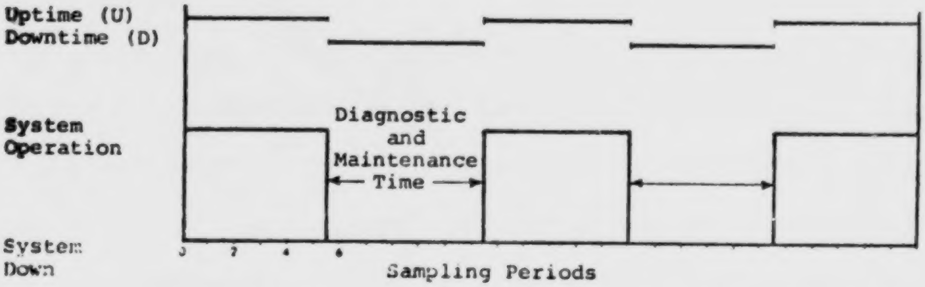
Example of Positive Biassed (Inaccurate) but Precise Measurements


Precision (σ)
True Value and Measured Average

Example of Accurate (No Bias) but Imprecise Measurements


Precision (σ)
True Value and Measured Average

Example of Precise and Accurate Measurements


Uptime (U)
Downtime (D)
System Operation
Diagnostic and Maintenance Time
System Down
Sampling Periods

Example Indicating a Measure of Completeness of Data U/(D + U)

Figure 1.4.2.1 Examples of data with varying degrees of precision, accuracy, and completeness

precise and accurate. Figure 1.4.2.1 also shows an example of completeness of data.

Each laboratory should have quantitative objectives set forth for each monitoring system in terms of completeness, precision, and accuracy of data. An example is included below for continuous measurement of carbon monoxide (non-dispersive infrared spectrometry) to illustrate the point.

1. Completeness - For continuous measurements, 75 percent or more of the total possible number of observations must be present. A summary of a minimum number of observations by time intervals (i.e., yearly, monthly, and hourly) is shown in Figure 1.4.17.1 (Data Validation).[3]

2. Precision - Determined with calibration gases, precision is ±0.5 percent full scale in the 0 through 58 mg/m^3 range.[4]

3. Accuracy - Depends on instrument linearity and the absolute concentrations of the calibration gases. An accuracy of ±1 percent of full scale in the 0 through 58 mg/m^3 range can be obtained.[4]

For further discussion of completeness, precision, accuracy and comparability, see the following:

1. Completeness and comparability, Section 1.4.17.
2. Precision and accuracy, Appendix G.

Employment of the elements of quality assurance discussed in Section 1.4. should lead to the production of data that are complete, accurate, precise, and comparable.

1.4.2.3 REFERENCES

1. Juran, J.M., (ed.). Quality Control Handbook. 2nd Ed. McGraw-Hill, New York, 1962. Sec. 2, pp. 4-8.

2. Feigenbaum, A.V. Total Quality Control. McGraw-Hill, New York, 1961. pp. 20-21.

3. Juran, J.M., and Gryna, F.M. Quality Planning and Analysis. McGraw-Hill, New York, 1970. pp. 375-377.

4. Nehls, G.J., and Akland, G.G. Procedures for Handling Aerometric Data. Journal of the Air Pollution Control Association, 23 (3):180-184, March 1973.

5. National Primary and Secondary Ambient Air Quality Standards. Federal Register, 36(84): 8194-8195, April 30, 1971. (Note: Copy available in Volume II of this Quality Assurance Handbook.)

1.4.3 ORGANIZATION

1.4.3.1 ABSTRACT

1. Organizing a quality assurance function includes establishing objectives, determining the amount of emphasis to place on each quality assurance element, preparing a quality assurance plan, identifying quality assurance problems to be resolved, and implementing the quality assurance plan.

2. Quality assurance is normally a separate function in the organization.

3. Quality assurance has input into many functions of an air pollution control agency. (See Figure 1.4.3.2 for details.)

4. The basic tools for quality assurance implementation are:

   a. Organization chart

   b. Job description. (See Figure 1.4.3.3 for job description for the Quality Assurance Coordinator.)

   c. Quality assurance plan

1.4.3.2 DISCUSSION

Organizing the quality assurance function[1] - Because of the differences in size, workloads, expertise, and experience in quality assurance activities among agencies adopting the use of a quality assurance system, it is useful here to outline the steps for planning an efficient quality assurance system.

1. Establish quality assurance objectives (precision, accuracy, and completeness) for each measurement system (Section 1.4.2).

2. Determine the quality assurance elements appropriate for the agency (Figure 1.4.).

3. Prepare a quality assurance plan (normally on a project basis) for all measurement systems (Section 1.4.23).

4. On the basis of the quality assurance plan, identify quality assurance problems which must be resolved.

5. Implement the quality assurance plan.

Location of the quality assurance function in the organization[2] - If practical, one individual within an organization should be designated the Quality Assurance (QA) Coordinator. The QA Coordinator should undertake activities such as quality planning, auditing, and reliability. The QA Coordinator should also have the responsibility for coordinating all quality assurance activity so that complete integration of the quality assurance system is achieved. The QA Coordinator should, therefore, gain the cooperation of other responsible heads of the organization with regard to quality assurance matters.

As a general rule, it is not good practice for the quality assurance function to be directly located in the organization responsible for conducting measurement programs. This arrangement could be workable, however, if the person in charge maintains an objective viewpoint.

Relationship of the quality assurance function to other functions - The functions performed by a comprehensive air pollution control program at the state or local level are shown in Figure 1.4.3.1.[3] The relationship of the quality assurance function to the other agency functions is shown in Figure 1.4.3.2. The role of quality assurance can be grouped into two categories:

1. Recommend quality assurance policy and assist its formulation with regard to agency policy, administrative support (contracts and procurements), and staff training.

2. Provide quality assurance guidance and assistance for monitoring networks, laboratory operations, data reduction, instrument maintenance and calibration, litigation, source testing, and promulgation of control regulations.

Basic tools for quality assurance implementation are:

1. The organization chart[4] - The quality assurance organization chart should display line and staff relationships, and lines of authority and responsibility. The lines of authority and responsibility, flowing from the top to bottom, are usually solid, while staff advisory relationships are depicted by dashed lines.

2. The job description[5] - The job description lists the responsibilities, duties, and authorities of the job and relationships to other positions, individuals, or groups. A sample job description for a Quality Assurance Coordinator is shown in Figure 1.4.3.3.

Management Services

- Agency policy
- Administrative and clerical support
- Public information and community relations
- Intergovernmental relations
- Legal counsel
- Systems analysis, development of strategies, long-range planning
- Staff training and development

Technical Services

- Laboratory operations
- Operation of monitoring network
- Data reduction
- Special field studies
- Instrument maintenance and calibration

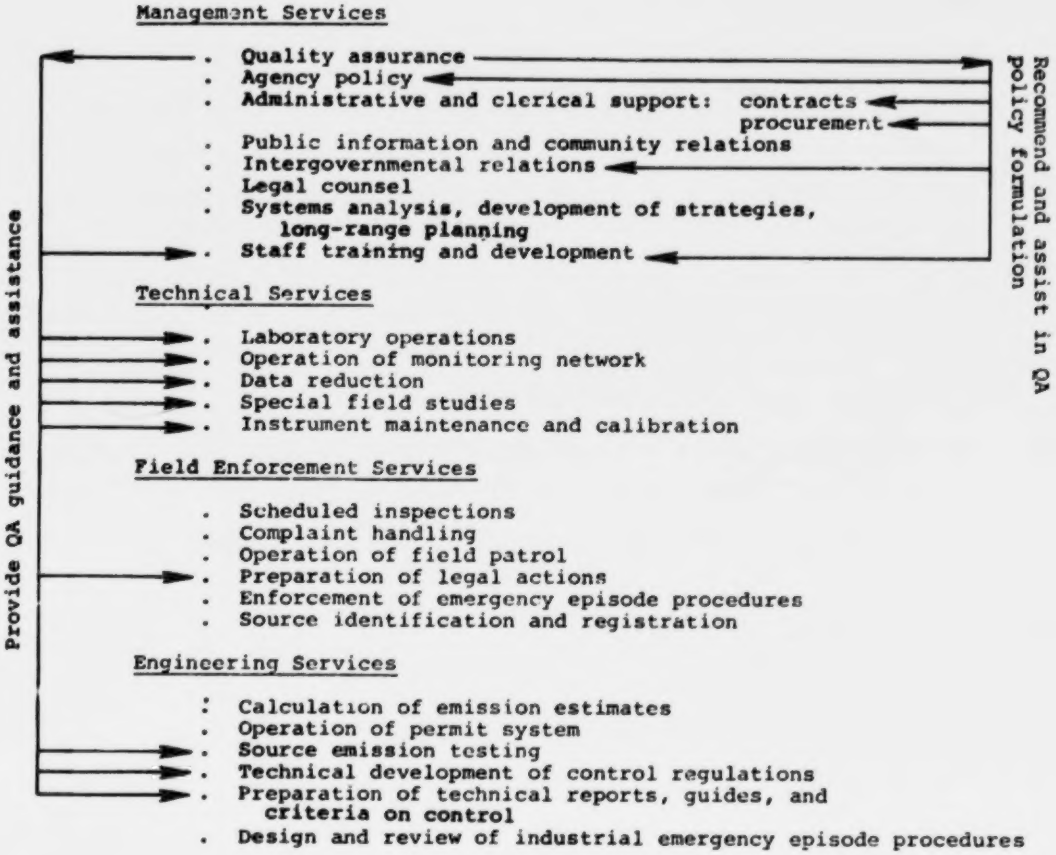
Field Enforcement Services

- Scheduled inspections
- Complaint handling
- Operation of field patrol
- Preparation for legal actions
- Enforcement of emergency episode procedures
- Source identification and registration

Engineering Services

- Calculation of emission estimates
- Operation of permit system
- Source emission testing
- Technical development of control regulations
- Preparation of technical reports, guides, and criteria on control
- Design and review of industrial emergency episode procedures

Figure 1.4.3.1 List of functions performed by comprehensive air pollution control programs


Provide QA guidance and assistance
Recommend and assist in QA policy formulation

Management Services

- Quality assurance
- Agency policy
- Administrative and clerical support: contracts
  procurement
- Public information and community relations
- Intergovernmental relations
- Legal counsel
- Systems analysis, development of strategies, long-range planning
- Staff training and development

Technical Services

- Laboratory operations
- Operation of monitoring network
- Data reduction
- Special field studies
- Instrument maintenance and calibration

Field Enforcement Services

- Scheduled inspections
- Complaint handling
- Operation of field patrol
- Preparation of legal actions
- Enforcement of emergency episode procedures
- Source identification and registration

Engineering Services

- Calculation of emission estimates
- Operation of permit system
- Source emission testing
- Technical development of control regulations
- Preparation of technical reports, guides, and criteria on control
- Design and review of industrial emergency episode procedures

Figure 1.4.3.2 Relationship of the quality assurance function to other air pollution control program functions

TITLE: Quality Assurance Coordinator

Basic Function

The Quality Assurance Coordinator is responsible for the conduct of the quality assurance program and for taking or recommending measures.

Responsibilities and Authority

1. Develops and carries out quality control programs, including statistical procedures and techniques, which will help agencies meet authorized quality standards at minimum cost.
2. Monitors quality assurance activities of the agency to determine conformance with policy and procedures and with sound practice; and makes appropriate recommendations for correction and improvement as may be necessary.
3. Seeks out and evaluates new ideas and current developments in the field of quality assurance and recommends means for their application wherever advisable.
4. Advises management in reviewing technology, methods, and equipment, with respect to quality assurance aspects.
5. Coordinates schedules for measurement system functional check calibrations, and other checking procedures.
6. Evaluates data quality and maintains records on related quality control charts, calibration records, and other pertinent information.
7. Coordinates and/or conducts quality-problem investigations.

Figure 1.4.3.3 Job description for the Quality Assurance Coordinator

3. The quality assurance plan - To implement quality assurance in a logical manner and identify problem areas, quality assurance plans are needed. These are normally prepared on a project basis. For details on preparation of quality assurance plans, see Section 1.4.23.

1.4.3.3 REFERENCES

1. Feigenbaum, A.V. Total Quality Control. McGraw-Hill, New York. 1961. Chapter 4, pp. 43-82.

2. Covino, C.P., and Meghri, A.W. Quality Assurance Manual. Industrial Press, Inc., New York. 1967. Step 1, pp. 1-2.

3. Walsh, G.W., and von Lehmden, D.J. Estimating Manpower Needs of Air Pollution Control Agencies. Presented at the Annual Meeting of the Air Pollution Control Association, Paper 70-92, June 1970.

4. Juran, J.M., (ed.). Quality Control Handbook, 2nd Edition. McGraw-Hill, New York. 1962. Section 6, pp. 242.

5. Industrial Hygiene Service Laboratory Quality Control Manual. Technical Report No. 78, National Institute for Occupational Safety and Health, Cincinnati, Ohio. 1974.

BIBLIOGRAPHY

1. Brown, F.R. Management: Concepts and Practice. Industrial College of the Armed Forces, Washington, D.C. 1967. Chapter II, pp. 13-34.

1.4.4 QUALITY PLANNING

1.4.4.1 ABSTRACT

1. Planning is thinking out in advance the sequence of actions needed to accomplish a proposed objective and communicating to the person or persons expected to execute these actions. Quality planning for air pollution measurements is designed to deliver acceptable quality data at a reasonable quality cost. Acceptable quality data is defined in terms of accuracy, precision, and completeness.

2. The critical characteristics in the total measurement system must be identified and controlled. These critical characteristics may be located in any one or all of the following spots:

   a. Sample collection.
   b. Sample analysis.
   c. Data processing.
   d. Associated equipment and analyzers.
   e. Users, namely operators and analysts.

3. Functional analysis (primarily ruggedness tests) and collaborative testing have been completed for many air pollution measurement methods. As a guide for the user of this Handbook, results from these studies are available in:

   a. A tabulated format in the method activity matrices for each method shown in Volume II, Ambient-Air Specific Methods and Volume III, Stationary-Source Specific Methods.

b. Collaborative study reports available from EPA are listed in Figure 1.4.15.1.

1.4.4.2 DISCUSSION

Approach to planning - The act of planning is thinking out in advance the sequence of actions necessary to accomplish certain objectives. In order that the planner may communicate his plan to the person or persons expected to execute it, the plan is written out with necessary criteria, diagrams, tables, etc.

Planning in the field of quality assurance for air pollution measurements must, of course, fundamentally be geared to deliver acceptable quality data at a reasonable quality cost. This objective is realized only by carefully planning many individual elements that relate properly to each other. These elements which make up the Quality Assurance Wheel shown in Figure 1.4.1, are discussed in Section 1.4 of this volume of the Handbook.

Specifications for data quality - The quality of data considered acceptable must be defined as quantitatively as possible. The three basic measures of quality of air pollution data are accuracy, precision, and completeness. Acceptance limits should be established for the completeness, precision, and accuracy of data recorded in the data bank, thus yielding acceptable levels of estimates of validity for all reportable data.

Acceptance limits for accuracy and precision of data are measurement method specific. Recommended acceptance limits, when available, are shown in the section entitled "Reference Method" for each method in Volume II, Ambient-Air Specific Methods and Volume III, Stationary-Source Specific Methods. In addition, accuracy and precision limits obtained by collaborative testing are available in EPA collaborative study publications. A list of available collaborative study publications on ambient air and stationary-source specific methods is shown in Figure 1.4.15.1.

Acceptance limits on criteria for completeness of data are recommended in Section 1.4.17, Data Validation.

Identification of critical characteristics - In the application of quality assurance measures, the total measurement system may be viewed as a complex system consisting of (1) the sample collection, (2) sample analyses, (3) data processing and the associated test equipment or analyzers and (4) the operators and analysts. The critical characteristics of this complex system are identified by functional analysis of which ruggedness testing* is one form. Ruggedness testing and collaborative study reports completed by the U.S. Environmental Protection Agency are listed in

---

* A series of empirical tests performed to determine the sensitivity (hopefully to confirm the insensitivity) of a measurement system (sample collection, sample analysis and data processing and the associated test equipment and analyzers and users.)

Figure 1.4.15.1. In addition, the method activity matrix in each measurement method in Volumes II and III is a tabulation of the most important characteristics (not necessarily all critical) for the measurement method for which control may be required.

Development of a QA plan[1] - The next step in the quality planning sequence is to determine which quality assurance element shown on the Quality Assurance Wheel (Figure 1.4.1) should be included as part of a quality assurance plan (QA plan). This is done by analyzing answers to questions posed to gain an understanding of the situation and needs. This analysis culminates in the selection of the most productive steps leading toward accomplishment of the objectives defined in terms of accuracy, precision, and completeness of data. Such questions might be:

1. What characteristics should be considered?
2. Which of these characteristics are most critical?
3. What acceptance limits should be assigned to the characteristics, particularly the most critical ones?
4. How often should these characteristics be checked?
5. What methods of measurement should be used to check the characteristics?
6. What action should be taken if the acceptance limits for characteristics are not met?

The answers to many of these questions are the intended purpose of the method activity matrices, which are a part of each method in Volumes II and III.

The finale of the quality planning should be a written document which includes the most important information the planner (normally a project officer or program manager in air pollution research and monitoring) thinks should be communicated to the person or persons executing the plan. This is called the QA plan. The recommended minimum content for a QA plan is discussed in Section 1.4.23 in the subsection entitled "Quality assurance plan format." The QA plan serves three main functions:

1. Most importantly, it is the culmination of a planning cycle, the purpose of which is to design into a project or program necessary provisions to assure quality data in terms of accuracy, precision, and completeness.

2. A historical record that documents the project (or program) plans in terms of (1) measurement methods used, (2) calibration standards and frequencies planned, (3) auditing planned, etc.

3. Provides the project officer or program manager a document that can be used to assess whether the quality assurance activities planned are being implemented and their importance for accomplishing the goal of quality data in terms of accuracy, precision, and completeness.

1.4.4.3 REFERENCES

1. Feigenbaum, A.V. Total Quality Control. McGraw-Hill, New York. 1961. pp. 134-149.

BIBLIOGRAPHY

1. Sindelar, F.J. Management Planning and Controls for an Effective Quality Function. Industrial Quality Control XVIII, 3:28-29. September 1961.

2. Ewing, David W., (ed.). Long-Range Planning for Management. Harper and Row, New York. 1964.

3. LeBreton, P.P., and Henning, D.A. Planning Theory. Prentice-Hall, Englewood Cliffs, New Jersey. 1961.

4. Steiner, G.A., (ed.). Managerial Long-Range Planning. McGraw-Hill, New York. 1963.

1.4.5 TRAINING

1.4.5.1 ABSTRACT

All personnel involved in any function affecting data quality (sample collection, analysis, data reduction, and quality assurance) should have sufficient training in their appointed jobs to contribute to the reporting of complete and high quality data.

The training methods commonly used in the air pollution control field are the following:

1. On-the-job training (OJT).
2. Short-term course training (normally 2 weeks or less).
3. Long-term course training (quarter or semester in length).

A list of recommended short-term course training is shown in Figure 1.4.5.1.

Training should be evaluated in terms of the trainee and the training per se. The following are techniques commonly used in the air pollution control field to evaluate training.

1. Testing (pre-training and post-training tests).
2. Proficiency checks.
3. Interviews (written or oral with the trainee's supervisor and/or trainee).

1.4.5.2 DISCUSSION

All personnel involved in any function affecting data quality (sample collection, analysis, data reduction, and quality assurance) should have sufficient training in their appointed jobs to contribute to the reporting of complete and high quality data. The Quality Assurance Coordinator should be concerned that the required training is available for these personnel and, when it is not, should recommend to management that appropriate training be made available.

Training objective[1,2] - The training objective should be to develop personnel to the necessary level of knowledge and skill required for air pollution measurement systems (ambient air and source emissions).

Training methods and availability - Several methods of training are available to promote achievement of the desired level of knowledge and skill required. The following are the training methods most commonly used in the air pollution control field; a listing of available training courses for 1975-1976 is given in Figure 1.4.5.1.

1. On-the-job training (OJT) - An effective OJT program could consist of the following:

   a. Observe experienced operator perform the different tasks in the measurement process.

   b. Study the operational portion of the method as described in the pollutant-specific portion of this Handbook, and use it as a guide for performing the operations.

Course Title	Length (days)	Location	Contact
Quality Assurance Training			
1. Quality Assurance for Air Pollution Measurement Systems (470)	3	RTP(1)	APTI(1)
2. Quality Management Institute	5	Storrs, CT	U of CT(2)
3. Managing for Quality	5	Milwaukee, WI	ASQC(3)
4. Applied Quality Assurance Engineering Operations	10	Los Angeles, CA	UCLA(4)
5. Quality Engineering	5	Milwaukee, WI	ASQC
6. Quality Audit - Development and Administration	3	Milwaukee, WI	ASQC
Air Pollution Measurement Method Training			
1. Source Sampling for Particulate Pollutants (450)	4	RTP	APTI
2. Source Sampling and Analysis for Gaseous Pollutants (468)	4.5	RTP	APTI
3. Visible Emission Certification (471)	3	Field(1)	APTI
4. Meteorological Instrumentation in Air Pollution (447)	5	RTP	APTI
5. Atmospheric Sampling (435)	4	Field	APTI
6. Reference Method for the Determination of Sulfur Dioxide in the Atmosphere - Pararosaniline Method (467A)	3	RTP	APTI
7. Reference Method for the Determination of Photochemical Oxidants Corrected for Interferences due to Nitrogen Oxides and Sulfur Dioxide (467C)	3	RTP	APTI
8. Analytical Methods for Air Quality Standards (464)	10	RTP	APTI
9. Measurement of Atmospheric Metals (436)	5	RTP	APTI
10. Gas Chromatographic Analysis of Air Pollutants (429)	10	RTP	APTI
11. Air Pollution Microscopy (420)	4.5	RTP	APTI

Figure 1.4.5.1 Selected quality assurance and air pollution training available in 1975-1976

Course Title	Length (days)	Location	Contact
Air Pollution Measurement Systems Training			
1. Air Quality Monitoring Systems - Planning and Administrative Concepts (463)	5	RTP	APTI
2. Site Selection for Ambient Air Monitoring (469)	3	RTP	APTI
3. Air Pollution Meteorology (411)	5	Field	APTI
4. Diffusion of Air Pollution - Theory and Application (423)	5	RTP	APTI
5. Air Quality and Emission Data Systems - AEROS (472)	3	Field	APTI
6. Statistical Evaluation of Air Pollution Data (426)	4	Field	APTI

Additional information may be obtained from:

(1) Registrar, Air Pollution Training Institute, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711 Note: Courses are offered at the Research Triangle Park (RTP) and throughout the USA (field).

(2) Dr. A. L. Jeffers, Box U-56, The University of Connecticut, Storrs, Connecticut 06268

(3) Registrar, Education and Training Institute, American Society for Quality Control, 161 West Wisconsin Avenue, Milwaukee, Wisconsin 53203

(4) P.O. Box 24901, Dept. K. UCLA Extension, Los Angeles, California 90024

Figure 1.4.5.1 (continued). Selected quality assurance and air pollution training available in 1975-1976.

c. Perform operation under the direct supervision of an experienced operator.

d. Perform operations independently but with a high level of quality assurance checks, utilizing the evaluation technique described later in this section to encourage high quality work.

2. Short-term course training - A number of short-term courses (normally 2 weeks or less) are available that provide knowledge and skills to implement more effectively an air pollution measurement system. Some of the courses are on the measurement methods *per se* and others provide training useful in the design and operation of the total or selected portions of the measurement system.

3. Long-term course training - Numerous universities, colleges, and technical schools provide long-term (quarter and semester length) academic courses in statistics, analytical chemistry, and other disciplines. The agency's training or personnel officer should be contacted for information on the availability of long-term course training.

Training evaluation - Training should be evaluated in terms of (1) level of knowledge and skill achieved by the operator from the training; and (2) the overall effectiveness of the training, including determination of training areas that need improvement. If a quantitative performance rating can be made on the operator during the training period (in terms of knowledge and skill achieved), this rating may also provide an assessment of the overall effectiveness of the training as well.

Several techniques are available for evaluating the operator and the training _per se_. One or more of these techniques should be used during the evaluation. The most common types of evaluation techniques applicable to training in air pollution measurement systems are the following:

1. _Testing_ - A written test before (pretest) and one after (post-test) training are commonly used in short-term course training. This allows the trainee to see areas of personal improvement and provides the instructor with information on training areas that need improvement.

2. _Proficiency checks_ - A good means of measuring skill improvement in both OJT and short-term course training is to assign the trainee a work task. Accuracy and/or completeness are commonly the indicators used to score the trainee's proficiency. The work tasks could be of the following form:

   a. Sample collection - Trainee would be asked to list all steps involved in sample collection for a hypothetical case. In addition, the trainee could be asked to perform selected calculations. Proficiency could be judged in terms of completeness and accuracy.

b. Analysis - Trainee could be provided unknown samples for analysis. As defined here, an unknown is a sample whose concentration is known to the work supervisor (OJT) or training instructor (short-term course training) but unknown to the trainee. Proficiency could be judged in terms of accuracy.

c. Data reduction - Trainees responsible for data reduction could be provided data sets to validate. Proficiency could be judged in terms of completeness and accuracy.

If proficiency checks are planned on a recurring basis, a quality control or other type chart may be used to show progress during the training period as well as after the training has been completed. Recurring proficiency checks are a useful technique for determining if additional training may be required.

3. Interviews - In some cases, a written or oral interview with the trainee's supervisor and/or trainee is used to determine if the training was effective. This interview is normally not conducted until the trainee has returned to the job and has had an opportunity to use the training. This technique is most often used to appraise the effectiveness of a training program (OJT or short-term course) rather than the performance of the trainee.

1.4.5.3 REFERENCES

1. Feigenbaum, A.V. *Total Quality Control*. McGraw-Hill, New York. 1961. pp. 605-615.

2. Feigenbaum, A.V. Company Education in the Quality Problem. *Industrial Quality Control*, *X*(6):24-29, May 1974.

BIBLIOGRAPHY

1. Juran, J.M., (ed.). *Quality Control Handbook*. 2nd edition. , McGraw-Hill, New York, 1962. Section 7, pp. 13-20.

2. Reynolds, E.A. Industrial Training of Quality Engineers and Supervisors. *Industrial Quality Control*, *X*(6):29-32, May 1954.

3. *Industrial Quality Control*, *23*(12), June 1967. (All articles deal with education and training.)

4. Seder, L.A. QC Training for Non-Quality Personnel. *Quality Progress*, *VII*(7):9.

5. Reynolds, E.A. Training QC Engineers and Managers. *Quality Progress*, *III*(4):20-21, April 1970.

1.4.6 PRETEST PREPARATION

1.4.6.1 ABSTRACT

1. A common practice in both ambient air monitoring and stationary source emission monitoring includes pretest preparation for a new project. The proper selection of sampling sites is fundamental in providing monitoring data that are both high quality and representative.

2. The pretest activities most important in ambient air monitoring system design are:

   a. Monitoring network size.

   b. Sampling station location.

3. The pretest activities most important in stationary source emission monitoring system design are:

   a. Process design and operation familiarity.

   b. Measurements performed to gather data for design of the sample collection program.

   c. Monitoring of process to determine representative conditions of operation.

1.4.6 2 DISCUSSION

A common practice in both ambient air monitoring and stationary source emission monitoring includes pretest preparation for a new project. During the pretest preparation, an on-site visit is normally conducted in order to complete administrative details for sample collection and to gather technical information for monitoring system design.

Ambient air monitoring system design - Factors that should be considered during ambient air monitoring system design are discussed in Volume II. The items most important during the pretest preparation are summarized here.

1. Monitoring network size. The design of the monitoring network depends on the objective of the project. The objective is normally one of the following: compliance monitoring, emergency episode monitoring, trend monitoring, or research monitoring. In considering the location of the network, one or more of the following will be important considerations: monitoring must be pollution oriented; monitoring must be population oriented; monitoring must be source oriented; and/or monitoring must provide area-wide representation of air quality.

Criteria should be provided for new project monitoring network design. By way of example, criteria for design of monitoring required by State Implementation Plans (SIP) are shown in Figure 1.4.6.1 for mechanical samplers and continuous samplers with regard to the distribution of stations in industrial, residential, and nonurban areas.[1] Figure 1.4.6.2 shows the requirements in the Federal Register for SIP monitoring by pollutant type.[2]

2. Sampling station location. The location of sampling stations within a monitoring network is influenced primarily by meteorological and topographic restraints. Meteorology (wind direction and speed) not only affect the geographical location of the sampling station but also the height of the sampling probe or sampler. Topographic features that have the greatest influence on final sampling

DISTRIBUTION OF MECHANICAL (INTEGRATED) SAMPLING STATIONS AS A FUNCTION OF TOTAL STATIONS IN A NETWORK[a]

Total number of stations	Number of stations in: Center city/ industrial	Residential zones	Nonurban
1	1	-	-
2	1	1	-
3	2	1	-
4	2	2	-
5	2	2	1
10	5	3	2
15	8	5	2
20	12	6	2
25	14	8	3
30	17	10	3

[a] Includes Hi-Vol sampler and SO_2, NO_2, and oxidant (ozone) collecting devices.

DISTRIBUTION OF AUTOMATIC (CONTINUOUS) SAMPLING STATIONS AS A FUNCTION OF TOTAL STATIONS IN A NETWORK[a]

Total number of stations	Number of stations in: Center city/ industrial	Residential zones	Nonurban[b]
1	1	-	-
2	1	1	-
3	2	1	-
4	2-3	1-2	-
5	3	2	-
6	4	2	-
10	6	4	-
15	10	5	-

[a] Includes SO_2, CO, HC, NO, NO_2, and oxidant (ozone).

[b] Where ozone damage has been identified in nonurban areas, surveillance may be necessary.

Figure 1.4.6.1 Recommended SIP monitoring station distribution.

Classification of region	Pollutant	Measurement methods or principle[i]	Minimum frequency of sampling	Region population	Minimum number of air quality monitoring sites[h]
I	Suspended particulates	High volume sampler	One 24-hour sample every 6 days[a]	Less than 100,000	4.
				100,000-1,000,000	4+0.6 per 100,000 population.[c]
				1,000,001-5,000,000	7.5+0.25 per 100,000 population.[c][e]
				Above 5,000,000	12+0.16 per 100,000 population.
		Tape sampler	One sample every 2 hours		One per 250,000 population[c] up to eight sites.
	Sulfur dioxide	Pararosaniline or equivalent[d].	One 24-hour sample every 6 days (gas bubbler).[a]	Less than 100,000	3.
				100,000-1,000,000	2.5+0.5 per 100,000 population.[c]
				1,000,001-5,000,000	6+0.15 per 100,000 population.[c][e]
				Above 5,000,000	11+0.05 per 100,000 population.
			Continuous	Less than 100,000	1.
				100,000-5,000,000	1+0.15 per 100,000 population.[c]
				Above 5,000,000	6+0.05 per 100,000 population.[c]
	Carbon monoxide	Nondispersive infrared or equivalent.[e]	Continuous	Less than 100,000	1.
				100,000-5,000,000	1+0.15 per 100,000 population.[c]
				Above 5,000,000	6+0.05 per 100,000 population.[c]
	Photochemical oxidants	Gas phase chemiluminesence or equivalent.[f]	Continuous	Less than 100,000	1.
				100,000-5,000,000	1+0.15 per 100,000 population.[c]
				Above 5,000,000	6+0.05 per 100,000 population.[c]
	Nitrogen dioxide	24-hour sampling method (Jacobs-Hochheiser method).	One 24-hour sample every 14 days (gas bubbler).[b]	Less than 100,000	3.
				100,000-1,000,000	4+0.6 per 100,000 population.[c]
				Above 1,000,000	10.
II	Suspended particulates	High volume sampler	One 24-hour sample every 6 days[a]		3.
		Tape sampler	One sample every 2 hours		1.
	Sulfur dioxide	Pararosaniline or equivalent[d].	One 24-hour sample every 6 days (gas bubbler).[a]		3.
			Continuous		1.
III[g]	Suspended particulates	High volume sampler	One 24-hour sample every 6 days[a]		1.
	Sulfur dioxide	Pararosaniline or equivalent[d].	One 24-hour sample every 6 days (gas bubbler).[a]		1.

[a] Equivalent to 61 random samples per year.

[b] Equivalent to 26 random samples per year.

[c] Total population of a region. When required number of samplers includes a fraction, round-off to nearest whole number.

d, e, and f [Reserved, 40 FR 7042, February 18, 1975]

[g] It is assumed that the Federal motor vehicle emission standards will achieve and maintain the national standards for carbon monoxide, nitrogen dioxide, and photochemical oxidants; therefore, no monitoring sites are required for these pollutants.

[h] In interstate regions, the number of sites required should be prorated to each State on a population basis.

[i] Named methods and principles, except the tape sampler method, are described in Part 50 of this chapter. The tape sampler method is described in Hemeon, W. C. L., Haines, G. F., Jr, and Ide, H. M., "Determination of Haze and Smoke Concentrations by Filter Paper Samplers", J. Air Pollution Control Association, Vol. 3, pp. 22-28, 1953. Use of these and other methods shall be as specified in § 51.17a.

Figure 1.4.6.2 Recommended SIP monitoring requirements by pollutant.

station location are physical obstructions in the immediate area that may alter air flows, e.g., trees, fences, and buildings; etc. Criteria should be provided for sampling station location for new projects.

Providing project criteria for network and station design before the on-site inspection is an important factor in the success of the project and in the quality and representativeness of the data.

Stationary source emission monitoring system design - Factors that should be considered during stationary source emission monitoring system design are discussed in Volume III. The items most important during the pretest preparation are:

1. Familiarization with process design and operation. The success of source emission monitoring requires familiarity with the process to be monitored. The following are areas in which familiarity is particularly important:

a. Process operation principle.

b. Process flow chart.

c. Variability of process operation in terms of flow rate of effluent to be monitored and concentration of pollutant in the effluent.

d. Process data that must be collected during sample collection, e.g., fuel burning rate.

e. Identify the key parameters and their representative levels of operation.

f. Sample collection site considerations, including sample site location (no turbulence due to upstream obstructions or bends), sample port access and size, size of platform for sample collection work, and utilities availability for sample collection equipment.

2. Measurements needed for design of the sample collection program - During the on-site visit, certain measurements are normally made that are required for the design of the sample collection program. These measurements include:

a. Dimensions of the stack or duct cross-section so that a sampling plan by equal cross-sectional areas can be determined.

b. Gas velocity, gas temperature, and gas moisture content so that requirements for isokinetic sampling can be calculated.

The selection of proper sampling sites and preliminary measurements required for monitoring system design during the pretest preparation is fundamental in providing monitoring data that are both high quality and representative.

1.4.6.3 REFERENCES

1. Air Quality Surveillance Network. Environmental Protection Agency, Office of Air Programs, Durham, N.C. Publication No. AP-98. May 1971.

2. Requirements for Preparation, Adoption, and Submittal of Implementation Plan. Federal Register, 36(158), August 14, 1971.

1.4.7 PREVENTIVE MAINTENANCE

1.4.7.1 ABSTRACT

1. The most important effect a good preventive maintenance program has is to increase measurement system reliability and thus increase data completeness.

2. Continuous pollutant analyzers commonly require daily service checks. An example of a daily checklist is shown in Figure 1.4.7.1. Daily service checks, preventive maintenance, and calibration should be integrated into a schedule. An example of a combined preventive maintenance-calibration schedule is shown in Figure 1.4.7.2.

1.4.7.2 DISCUSSION

Importance of preventive maintenance - As defined here, preventive maintenance is an orderly program of positive actions (equipment cleaning, lubricating, reconditioning, adjustment and/or testing) for preventing failure of monitoring systems or parts thereof during use. The most important effect a good preventive maintenance program has is to increase measurement system reliability and thus increase data completeness. Conversely, a poor preventive maintenance program will result in increased measurement system downtime (i.e., decrease in data completeness) and in increased unscheduled maintenance costs; and may cause distrust in the validity of the data. In ambient air monitoring, data completeness criteria are used to validate data. A complete discussion of data completeness may be found in Section 1.4.17 (Data Validation).

Preventive maintenance schedule - Project officers should prepare and implement a preventive maintenance schedule for measurement systems. The planning required to prepare the preventive maintenance schedule will have the effect of:

1. Highlighting that equipment or those parts thereof that are most likely to fail without proper preventive maintenance.

2. Defining a spare parts inventory that should be maintained to replace worn-out parts with a minimum of downtime.

A specific preventive maintenance schedule should relate to the purpose of testing, environmental influences, physical location of analyzers, and the level of operator skills. Checklists are commonly used to list specific maintenance tasks and frequency (time interval between maintenance).

Continuous pollutant analyzers commonly require daily service checks. By way of example, Figure 1.4.7.1 is a daily instrument service checklist and record used in a major EPA monitoring network for ambient air measurement of NO_2.[1] In this particular monitoring network, ambient air measurement data for NO_2 are recorded on strip charts and sent weekly to a central location for validation and data

CITY ______________________

SITE LOCATION ______________________

____

SITE NUMBER ______________________

AT LAST CALIBRATION	NO	NO_2
SCALE RANGE		
ZERO KNOB SETTING		
SPAN KNOB SETTING		

SERIAL NUMBER ______________________

DATE LAST CALIBRATION ______________________

DATE							
OXYGEN PRESSURE - INST.							
OXYGEN CYLINDER PRESSURE							
UNADJUSTED NO ZERO READING							
UNADJUSTED NO_2 ZERO READING							
NO ZERO KNOB SETTING (NEW)							
NO SPAN KNOB SETTING							
NO_2 ZERO KNOB SETTING (NEW)							
NO_2 SPAN KNOB SETTING							
✓ VALVE IN NO-NO_2-NO_x POSITION							
✓ NO-NO_2 RANGE IN 0.5 POSITION							
✓ INSPECT OXYGEN LINE							
✓ INSPECT INLET LINE, PROBE, & FILTER HOLDER							
VACUUM GAUGE READING							
COMMENTS OR PROBLEMS:							
OPERATOR INITIAL							

Figure 1.4.7.1 Daily checklist for NO_2 analyzer.

reduction. The instrument checklist shown in Figure 1.4.7.1 is sent with the strip charts. This checklist provides a record of daily service checks and, in addition, includes operations data used in the validation of the strip-chart data.

When a project includes several sensors (air pollution and/or meteorological), it becomes important to integrate checklists into a preventive maintenance schedule. Since calibration sensors are commonly the responsibility of the operator in addition to preventive maintenance, and since calibration tasks may be difficult to separate from preventive maintenance tasks, a combined preventive maintenance-calibration schedule is often advisable. An example of a combined preventive maintenance-calibration schedule for a major EPA ambient air monitoring project is shown in Figure 1.4.7.2.[2]

This combined schedule is read as follows: Numbers along the vertical (1-36) and horizontal (1-10) part of the schedule refer to the Julian date. The numbers in the boxes indicate the tasks to be completed according to Julian date. The following should be noted in this combined schedule:

1. The schedule is prepared by Julian date, not calendar date. For example, Julian date 10 is January 10 and Julian date 33 is February 2.

2. The operator is provided a reference to the project operations manual for the exact procedure to follow.

	1	2	3	4	5	6	7	8	9	10
0		1,7	5,14,27	12,13	2,38,40	6,52	23,26,28	3	6	31,32
1	33,34	4,49,50, 53	45,47,43	9,10,11, 28	1	7	5,27		2	8
2	23,26,28	3	6	21,22,30, 31	33,34	4,46,53	51,52	9,10,11, 28	1	7
3	5,27	55	2	3,49,50	23,26,28	3	6	31,32	33,34,35	4,53
4	45,47,48	9,10,11, 28	1	7	5,27		2	8,52	23,26,28	3
5	6	14,15,31	33,34,36	4,53	49,50	9,10,11, 12,13	1.28	7	5,27	
6	2,47	3	22,26,28	3	6	31,32	33,34	4,53	45,47,48	9,10,11, 52
7	1,28	7	5,27		2	8,21,22, 23,26,30	24,49,50	3,23,46		
8	33,34,35	4,53		9,10,11	1,18	7	5	27		
9	23,26	3,28,52	6	31,32	19,33,34				1,28	7
10	5	27	2	9	23,26				31,32,33 34	4
						27	2	8,52	23,26	3,28
						9,10,11, 21,22	1,28	7,46	5	27,51
34	2	8	23,26	3,28	6	33,34	31,32	4	45,47,48 53	9,10,11,33, 34,52
35	1,28	7	5,54	27	2	8,49,50	23,26	3,28	6	
36	31	4	53	25	9,10,11, 28					

TASK NO.	TASK	REFERENCE
1	Calibrate MF-1D, MF-2D, MF-3D, MF-2G, MF-1H and MF-2V	Paragraph 3.2.2
2	Calibrate MF-1G, MF-1V and MF-1C	Paragraph 3.2.2
3	Calibrate MF-2H, MF-1S and MF-4H	Paragraph 3.2.2
4	Calibrate MF-3H and MF-1P	Paragraph 3.2.2
5	Calibrate NO_2 analyzer (5-point)	Paragraph 3.2.1
6	Calibrate SO_2 analyzer (5-point)	Paragraph 3.2.1
7	Calibrate NO and NO_x analyzers (5-point)	Paragraph 3.2.1
8	Calibrate O_3 analyzer (5-point)	Paragraph 3.2.1
9	Inspect compressor and vacuum pumps shock mounts	Paragraph 3.2.9
10	Clean compressor intake filters	Paragraph 3.2.10
11	Change pump box filters	Paragraph 3.2.11
12	Clean vacuum panel filter	Paragraph 3.2.20
13	Replace vacuum pump filters	Paragraph 3.2.12
14	Replace vacuum pump vanes	Refer to vendors literature

(Note: MF means mass flow meter.)

Figure 1.4.7.2 Combined preventive maintenance-calibration schedule for a major EPA ambient-air monitoring project.

For example, the procedure for Task No. 6 (calibrate SO_2 analyzer) is contained in paragraph 3.2.1 of the project operations manual.

3. At a glance, the operator knows each day the scheduled calibration and preventive maintenance tasks that must be performed. For example, on Julian date 63, Task No. 22 (lubricate pump box blower motor), Task No. 26 (leak check H_2/O_2 generator), and Task No. 28 (replace analyzers sample filters) are all to be completed.

This EPA ambient air monitoring project involves continuous air pollution and meteorological measurements with intermediate computer storage of measurement data at the field sampling station and subsequent telemetering of measurement data to a central computer at the Research Triangle Park, North Carolina. Figure 1.4.7.3 is a list of preventive maintenance and calibration tasks that are performed at each field sampling station.[2] A partial list of these tasks was shown previously in Figure 1.4.7.2. This list of tasks is shown as an example of the scope of preventive and calibration tasks that may be required on a large project. In addition to the tasks shown in Figure 1.4.7.3, zero and span checks are performed periodically on each air pollution sensor, as well as daily service checks on each sensor.

Preventive maintenance records - A record of all preventive maintenance and daily service checks should be

TASK NO.	TASK	TASK NO.	TASK	TASK NO.	TASK
1	Calibrate MF-1D, MF-2D, MF-3D, MF-2G, MF-1M and MF-2V	20	Clean dryer relief valve	39	Lubricate air conditioner motors
2	Calibrate MF-1G, MF-1V and MF-1C	21	Replace compressor rings, valves and springs	40	Check air conditioner cabinet water drain
3	Calibrate MF-2H, MF-1S and MF-4H	22	Lubricate pump box blower motors	41	Clean air conditioner coil
4	Calibrate MF-3H and MF-1P	23	Replace H_2/O_2 generator water tank	42	Wax air conditioner cabinet
5	Calibrate NO_2 analyzer (5-point)	24	Service and adjust H_2/O_2 generator	43	Calibrate wind speed instrument
6	Calibrate SO_2 analyzer (5-point)	25	Change molecular sieves	44	Calibrate wind direction instrument
7	Calibrate NO and NO_x analyzers (5-point)	26	Leak check H_2/O_2 generator	45	Fill water bath
8	Calibrate O_3 analyzer (5-point)	27	Inspect sample lines	46	Change SO_2 permeation tube
9	Inspect compressor and vacuum pumps shock mounts	28	Replace analyzers sample filters	47	Check meteorological tower guy line tension
10	Clean compressor intake filters	29	Clean sample manifold	48	Inspect meteorological tower and instruments
11	Change pump box filters	30	Lubricate sample manifold blower motor	49	Replace particulate manifold motor
12	Clean vacuum panel filter	31	Leak check calibration system	50	Replace hi-vol motor
13	Replace vacuum pump filters	32	Leak check pressure panel	51	Replace NO-NO_x analyzer exhaust filter (charcoal)
14	Replace vacuum pump vanes	33	Change data tape	52	Calibrate ambient temperature sensor
15	Replace vacuum relief valve	34	Clean tape deck transport mechanism	53	Calibrate relative humidity sensor
16	Replace dryer ball check valve	35	Check tape deck skew and tape tracking	54	Calibrate xincom
17	Clean dryer solenoid valves	36	Check tape deck head wear	55	Weigh fire extinguishers
18	Clean dryer check valve	37	Replace tape deck reel motors and capstan motor		
19	Replace water trap automatic drain and filter	38	Clean air conditioner filters		

Figure 1.4.7.3 List of tasks for preventive maintenance and calibration for a major EPA ambient air monitoring project

maintained. Normally the daily service checklists shown in Figure 1.4.7.1 are stored with the measurement data. An acceptable practice to follow for recording completion of tasks listed in Figure 1.4.7.2 is to maintain a preventive maintenance-calibration multiple-copy record book. After tasks have been completed and entered in the record book, a replicate copy of each task is removed by the operator and sent to the supervisor for review and filing. The record book is stored at the sampling station for future reference.

1.4.7.3 REFERENCES

1. EPA, Quality Assurance and Environmental Monitoring Laboratory, Research Triangle Park, North Carolina 27711.

2. EPA, Human Studies Laboratory, Research Triangle Park, North Carolina 27711.

BIBLIOGRAPHY

1. EPA, Office of Air Programs. Field Operations Guide for Automatic Air Monitoring Equipment. Publication No. APTD-0736. Research Triangle Park, NC. 27711 1971.

2. Hubert, C.I. Preventive Maintenance of Electrical Equipment. McGraw-Hill, New York, N.Y. 1955. Part 1.

3. Spencer, J. Maintenance and Servicing of Electrical Instruments. Instruments Publishing Co., Inc., Pittsburgh, PA. 1951.

1.4.8 SAMPLE COLLECTION

1.4.8.1 ABSTRACT

1. Pre-sample and post-sample collection checks should be performed by the operator on the sample collection system. These checks are measurement-method specific. Recommended pre-sample and post-sample checks for each sample collection system are given for each method in Volume II, Ambient-Air Specific Methods, and Volume III, Stationary-Source Specific Methods.

2. Control checks should be performed during the sample collection. These checks are used to determine the performance of the sample collection system. Recommended control checks are given in each method in Volumes II and III.

3. Control charts should be used to record results from selected control checks to determine when the sample collection system is out-of-control and to record the corrective action taken.

1.4.8.2 DISCUSSION

Pre-sample and post-sample collection checks - Checks should be made by the operator prior to the actual sample collection. Pre-sample collection checks normally include:

1. A leak check on the sample collection system. Since the entire sample collection system is normally under a vacuum, it is very important to be sure that all parts of

the system are assembled so that air does not leak into the sample gas stream.

2. Specific checks on components of the sample collection system, such as the liquid level in bubblers.

The amount of post-sample collection checking is dependent on the sample collection technique. When samples are collected for subsequent laboratory analyses, selected post-sample checks on the sample collection system and the inspection of the sample per se are a common practice. When continuous air pollution sensors are used, sensor calibrations normally replace the post-sample collection checks. See Section 1.4.12 for more information on sensor calibration.

Pre-sample and post-sample collection checks are measurement-method specific. Recommended checks for the sample collection system are given in each method in Volumes II and III.

Control checks during sample collection - Control checks should be performed during the actual sample collection. These checks are used to determine the performance of the sample collection system.

Control charts - Results from selected control checks should be recorded on control charts. This will help the operator and supervisor to determine when the sample collection system is out-of-control and will provide a record of corrective action taken.

1.4.9 SAMPLE ANALYSIS

1.4.9.1 ABSTRACT

1. Function checks should be performed by the analyst to check the validity of the sample and performance of the equipment.

2. Control checks should be performed during the analyses. These checks are used to:

   a. determine the performance of the analytical system.

   b. quantitate the variability in results from the analytical system in terms of precision and accuracy.

3. Control charts should be used to record results from selected function and control checks to determine when the analytical system is out-of-control and record the corrective action taken.

1.4.9.2 DISCUSSION

Function checks - Function checks are performed to verify the stability and validity of the sample and the performance of the analytical equipment. The analyst should be provided with written performance specifications for each function check, accompanied by recommended action if the specifications are not met.

Control checks - Control checks should be performed during analysis. These checks are made on all analyses or intermittently after a specified number of analyses.

Some control checks are part of the routine analysis and are performed by the analyst to determine the performance of the analytical system. These control checks include the use of sample blanks to observe zero concentration drift; spiked samples to determine percentage of sample recovery during intermediate analysis extraction steps; and sample aliquots to observe within- and between-run variability for the entire analysis.

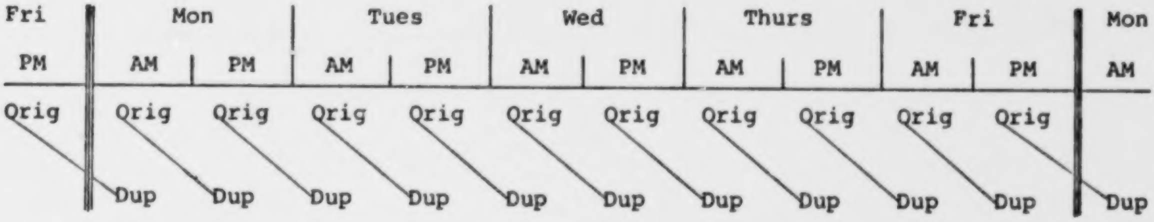
Other control checks are performed by the analyst intermittently to quantitate analysis variability in terms of precision and accuracy. Control samples normally used for this purpose are sample aliquots to determine precision and standard reference materials and standard reference samples to determine accuracy in addition to precision. A list of some available standard reference materials is shown in Figure 1.4.12.1.

Analysis variability must be explained in terms of possible sources of variability to make this measurement most meaningful. The three common ways to report precision, are: replicability, repeatability, and reproducibility. A more detailed discussion is included in Appendix A.

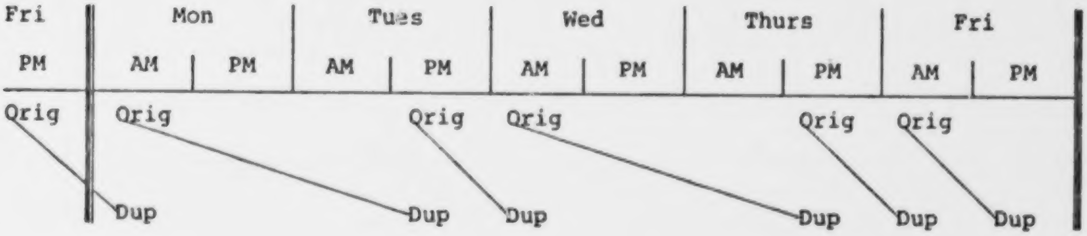
Source of variability	Replicability	Repeatability	Reproducibility
Sample	Same	Same	Same
Analyst	Same	{At least one of these must be different}	Different
Apparatus	Same		Different
Day	Same		Same or different
Laboratory	Same	Same	Different

A good scheme using environmental samples to determine repeatability is shown in Figure 1.4.9.1. The procedure involves duplicate analyses performed on a staggered time schedule, which allows for a better appraisal of laboratory variability than if the duplicate is analyzed immediately after the original sample. The number of duplicates removed for analysis is dependent on the number of analyses performed during the a.m. or p.m. time cycle shown in Figure 1.4.9.1. The number of duplicates removed for analysis would normally range from 1 to 5, with a minimum of one duplicate being analyzed for every a.m. or p.m. time span in which a sample set is analyzed. If the variability of the analyst is being monitored during routine analysis, efforts should be made to submit control samples in a manner so that the analyst does not give them special attention.

A. Analysis performed every day


Fri
Mon
Tues
Wed
Thurs
Fri
Mon
PM
AM
PM
AM
PM
AM
PM
AM
PM
AM
PM
AM
Orig
Dup

B. Analysis performed intermittently


Fri
Mon
Tues
Wed
Thurs
Fri
PM
AM
PM
AM
PM
AM
PM
AM
PM
AM
PM
Orig
Dup

Orig - analysis on original sample.

Dup - analysis on duplicate of original sample.

Figure 1.4.9.1 Analysis of duplicate samples

Control charts - Results from selected function checks and control samples should be recorded on control charts. This will allow the analyst and supervisor to know exactly when the analytical system is out-of-control, which part of the system is the probable cause, and when a corrective action is taken.

1.4.10 DATA REPORTING

1.4.10.1 ABSTRACT

1. Human error is the most common source of error in data reporting. However, individuals using measurement systems with data logging devices to automate data handling from continuous sensors should be concerned that only the sensor analog signal and not electronic interferences are converted during the digital readout.

2. Procedures for reviewing data at the operational, as well as the managerial levels, should be provided. Control charts are a common tool used to review data from critical characteristics in a measurement system.

3. A brief discussion of control charts is given in this section, with a more detailed discussion given in Appendix H, "Control Charts."

1.4.10.2 DISCUSSION

Source of data reporting errors - Measurements of the concentration of pollutants, either in the ambient atmosphere or in the emissions from stationary sources, are assumed to be representative of the conditions existing at the time of the sample collection. The extent to which this assumption is valid depends on the sources of error and bias inherent in the collection, handling, and analysis of the sample.

Besides the sampling and analytical error and bias, human error may be introduced any time between sample collection and sample reporting. Included among the human errors are such things as failure of the operator/analyst to record pertinent information, mistakes in reading an instrument, mistakes in calculating results, and mistakes in transposing data from one record form to another. Data handling systems involving the use of computers are susceptible to keypunching errors and errors involving careless handling of magnetic tapes and other storage media. Although it cannot be completely avoided, human error can be minimized.

Data reporting techniques and error sources depend on the type of sensor measurement system - Measurement sensors for pollutant concentration and meteorological conditions may be classified by their sample collection principle into two categories: (1) Integrated, and (2) Continuous. Pollutant measurement systems may be either integrated or continuous, whereas meteorological measurement systems are normally always continuous.

In the integrated sample collection principle, a discrete sample is collected in some medium and is normally sent to a laboratory for analysis. Both the field operator and the laboratory analyst can make errors in data handling.

In the continuous sample collection principle, an analytical sensor produces a direct and continuous readout of the pollutant concentration or meteorological parameter.

The readout may be a value punched on paper tape or recorded on magnetic tape. In addition, some continuous measurement systems may also use telemetry to transmit data to a data processing center. Both human and machine errors can occur in data handling in this type of system.

Data errors in integrated sampling - For ambient air monitoring, the operator records information before and after the sample collection period. For source emission testing, the operator records information during the sample collection period in addition to before and after it. Acceptance limits should be set for data pertaining to flow rates, etc., and the operator should invalidate or "flag" sampling data when values fall outside these limits. Many of these acceptance limits are included as part of the measurement method write-up. These and other recommended acceptance limits are tabulated in the method activity matrices for each method in Volume II, Ambient Air Specific Methods and Volume III, Stationary Source Specific Methods. Questionable measurement results may indicate the need for calibration or maintenance.

The analyst in the laboratory reads measurements from balances, colorimeters, spectrophotometers, and other instruments; and records the data on standard forms or in laboratory notebooks. Each time values are recorded, there is a potential for incorrectly entering results. Typical recording errors are transposition of digits (e.g., 216 might be incorrectly entered as 126) and incorrect decimal point location (e.g., 0.0635 might be entered as 0.635). These kinds of errors are difficult to detect. The supervisors must continually stress the importance of accuracy in recording results.

Acceptance limits contained in the measurement method write-up and those shown in the method activity matrix should be used by the analyst to invalidate or "flag" analysis data when values fall outside these limits.

Data errors in continuous sampling - Continuous air monitoring systems may involve either manual or automated data recording. Automated data recording may involve the use of a data logging device to record data on paper tape or magnetic tape at the remote sampling station, or the use of telemetry to transmit data on-line to a computer at a central facility.

Manual reduction of pollutant concentration data from strip charts can be a significant source of data errors. In addition to making those errors associated with recording data values on record forms, the individual who reads the chart can also err in determining the time-average value. Usually the reader estimates by inspection the average concentration. When the temporal variability in concentration is large, it is difficult to determine an average concentration. Two people reading the same chart may yield results that vary considerably.

Technicians responsible for reducing data from strip charts should be given training. After a technician is shown how to read a chart, his results should be compared with those of an experienced technician. Only after he has demonstrated the capability to obtain satisfactory results should a technican be assigned to a data reduction activity.

Periodically a senior technician should check strip charts read by each technician. Up to 10 percent of the data reported by each technician should be checked, depending on time availability and past history of error occurrence. If an individual is making gross errors, additional training may be necessary.

Because manual chart reading is a tedious operation, a drop in productivity and an increase in errors might be expected after a few hours. Ideally, an individual should be required to spend only a portion of a day at this task.

The use of a data logging device to automate data handling from a continuous sensor is not a strict guarantee against data recording errors. Internal validity checks are necessary to avoid serious data recording errors. There are two sources of error between the sensor and the recording medium: (1) the output signal from the sensor and (2) the errors in recording by the data logger.

The primary concern about the sensor output is to ensure that only the sensor analog signal and not electronic interferences be converted to a digital readout. Internal validity checks should be planned to "flag" spurious data resulting from electronic interferences.

For a system involving the use of telemetry, it is also necessary to include a validity check for data transmission.

Errors in computations - To minimize computational errors, operators and analysts should follow closely the formulae, calculation steps, and examples given for each method, using the calculation instructions and forms provided. As an example, Volume II provides instructions for calculations of air volume and air concentration for the Reference Method for the Determination of Suspended Particulates in the Atmosphere (High Volume Method).

Recommended audits on computations are shown in each method in Volumes II and III.

Control charts - Procedures for reviewing data at the operational as well as the managerial levels should be provided. Review of measurement results from control samples used during analysis, for example, can indicate out-of-control conditions that would yield invalid data from subsequent analyses, if the conditions are not corrected immediately. At the managerial level, periodic review of data can indicate trends or problems that need to be addressed to maintain the desired level of precision and accuracy. One common tool for statistical analysis of data at both the operational and the managerial levels is the control chart. The major steps in constructing the control chart are outlined here, and in greater detail in Appendix H.

The control chart provides a tool for identifying the systematic variation (assignable cause) from the system indeterminate variation (random). This technique displays data in a form that graphically compares the variability of all test results with the average or expected variability of small groups of data.

The steps to consider in the application of control charts are the following:

1. Select critical characteristics in the measurement system to audit. Measurement method characteristics are tabulated in the method activity matrix for each method in Volumes II and III. The critical characteristics are determined during quality planning (Section 1.4.4) and should be summarized in the quality assurance plan (Section 1.4.23).

2. If audit (reference) standards are used, obtain the necessary materials.

3. Select the data quality objective to audit:

a. Precision - Use the range ($\bar{R}$) deviation of the test results.

b. Accuracy - Use the average deviation of test results from the accepted value of reference standards (mean, $\bar{X}$).

4. Choose the audit size and frequency:[2]

a. Size - For analysis, air pollution analysts will often be dealing with samples of 2, which will form most subgroups.

b. Frequency of subgroup sampling - Changes are detected more rapidly as the sampling frequency is increased. Audit rates of 7 percent are recommended for many characteristics shown in the method activity matrices.

5. Set control limits. Control limits (CL) are normally set at 3 times the standard deviation for R or $\bar{X}$. Control limits can be calculated using the computational formulae and forms given in Appendix H of Volume I of this Handbook.

6. The control chart should be maintained either by the operator/analyst or the supervisor. The control chart should be kept up to date.

Recommended control chart types and control limits are given where appropriate in Volumes II and III. The control limits indicated are preliminary and are suggested as a point of departure when the laboratory has no data on which to base calculation of control limits. After establishing 15 to 20 data points (subgroup means and ranges), the control limits should be reestablished on the basis of these data. If the new control limits are narrower than those recommended in the method, the former should be used. After this initial calculation, control limits should be recalculated every 3 to 6 months, or whenever significant data trends or shifts become obvious.

The control chart is actually a graphical presentation of quality control effectiveness. A typical chart with indicated actions is given in Figure 1.4.10.1. If the procedure is "in control," the results will almost always fall within the established control limits. Further, the chart will disclose trends and cycles from assignable causes that can be corrected promptly.

A more complete discussion of how to choose the type of control chart, how to determine control limits, how to determine subgroup size, and how to interpret control charts for out-of-control conditions is shown in Appendix H.

1.4.10.3 REFERENCES

1. Control Chart Method of Controlling Quality During Production. ANSI Standard 21.3. November 1958. pp. 9 and 32.

BIBLIOGRAPHY

1. Kelley, W.D. Control Chart Techniques. Statistical Method-Evaluation and Quality Control for the Laboratory. August 1968. U.S. DHEW, Public Health Service, Cincinnati, Ohio. p. 3.

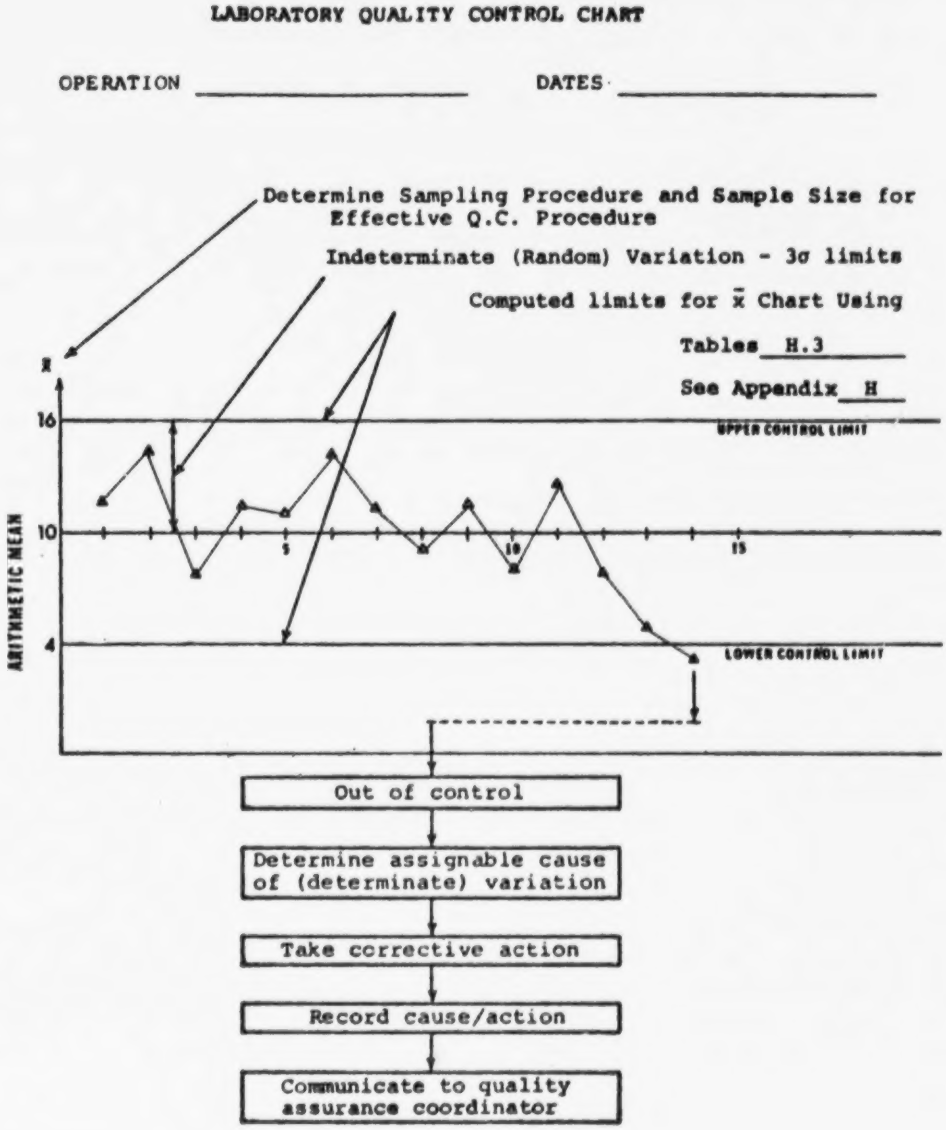


Figure 1.4.10.1 A typical quality control chart for means and indicated actions.

1.4.11 PROCUREMENT QUALITY CONTROL

1.4.11.1 ABSTRACT

1. Contracts for the procurement of (1) materials and equipment and (2) research and monitoring services should specify quality assurance requirements. Quality assurance requirements for research and monitoring services are summarized in Section 1.4.23 (Quality Assurance Plans for Projects and Programs). Quality assurance considerations for materials and equipment are included in this section.

2. Quality assurance requirements in contracts for materials and equipment most common to air pollution measurement systems are:

a. First article configuration inspection (FACI).

b. Systems acceptance test (SAT).

c. Certificate of analysis.

d. Reliability testing.

3. Recommended procurement specifications for materials and equipment used in air pollution measurement systems are listed in the method activity matrix for each method in Volume II (Ambient-Air Specific Methods) and Volume III (Stationary-Source Specific Methods) of this Handbook.

1.4.11.2 DISCUSSION

Contracts for the procurement of (1) materials and equipment and (2) research and monitoring services should specify quality assurance requirements. The quality assurance requirements specified for services should be consistent with the intended use of the measured data. During procurement of service contracts, each project officer should review Section 1.4.23 (Quality Assurance Plans for Projects and Programs) and determine which quality assurance elements are most important. The project officer should also require in the service contract that a quality assurance plan be submitted for review and approval by the project officer. In large service contracts, quality assurance information should be sought as part of the contract proposal received from bidders. This information is useful in evaluating the capability of providing quality data by each bidder. For U.S. government agencies, this request for quality assurance information should be included as part of the "Request for Proposal (RFP)." The RFP is an advertisement by the purchaser (a government agency) for proposals to complete the task described. An example of quality assurance information requested in an RFP for a major air pollution monitoring network for the U.S. Environmental Protection Agency is shown in Figure 1.4.11.1.

The remainder of this DISCUSSION deals with procurement quality control for materials and equipment.

Quality assurance requirements common to procurement - The quality assurance requirements most common to procurement of materials and equipment include the following:

1. The Contractor shall provide a full description of his proposed work team, giving as a minimum for each employee both full or part-time the following:

   a. Job title
   b. Skill level and type
   c. Brief description of duties
   d. Typical weekly deployment by location and time of day.

   For subcontractors, contractor shall state reason for subcontractor, particular capabilities, and expected role in system operation and maintenance. If one or more of subcontractor's personnel are considered to be critical, resumes shall be provided. Contractor shall provide resumes of its key personnel showing their experience and qualifications.

2. Contractor shall provide a plan of preventive maintenance of the remote stations including action to be undertaken and frequency.

3. Contractor shall further provide an estimate of mean time between failure for each major system component which its operation and maintenance plan shall achieve.

4. Contractor shall propose a quality assurance plan including quality control activities, calibrations, review and incorporation of results. Frequency of activities, traceability to reference materials (NBS preferable), etc. shall be included.

5. Contractor is encouraged to propose such changes in operating procedures, conditions, or software as it may deem advantageous to the government in terms of cost or enhancement of the monitoring system and related facilities. Such proposal shall provide a description of the change and the benefits thereof. If a proposal combines both equipment and procedural changes, for instance hardware and software, contractor may submit separate or combined proposals as deemed best for presentation of the modification.

6. Operation of the monitoring system during the period of this contract will result in the development of the most extensive data base on air quality to date. Concurrent with this operation is the development of much information on instrument and system performance. Incorporation of this information in assessing the degree of data base accuracy, developing extensive trouble-shooting guidelines, and providing feedback for system management is important.

Figure 1.4.11.1. Quality assurance information requested in an RFP for a major air pollution monitoring network.

Contractor shall develop a work plan achieving the goals above. Such plan shall state the formal procedures for review of system performance to incorporate the results of field examination of system components for the purpose of validating or invalidating data. The effect of the timeliness of processing data will be impacted and contractor shall consider such in its plan. Data shall be considered to be in one or three classes: valid, questionable, or invalid. Valid data meet all reasonable criteria of acceptance as accurate representations. The entire system objective is oriented toward the development of data for inclusion in the valid data class. Instrument failures or system malfunctions result in the loss of data, which are considered to be invalid for cause. Performance of an instrument beyond specification limits causes the data to be questionable. Field investigation of these instruments may result in invalidating data for cause. Contractor shall state how it shall incorporate field experience in validation activities. Further instruments may be discovered to have malfunctioned and retrospective action shall be taken to invalidate data previously considered valid. Procedural changes necessary to implement this activity shall be effected.

Operation and maintenance manuals provide a description of the system. Continuing revisions are necessary to incorporate the extensive experience gained from routine system operation. Contractor shall propose procedures it shall use to incorporate field experience into troubleshooting guidelines for manual revision. Contractor shall show flow diagrams of information resulting in consolidated manual revisions giving time intervals for each activity.

Contractor shall show in this plan how information of system performance is available for management review, both contractor's and EPA's management. This feedback is important in determining revisions of maintenance procedures and frequency.

Contractor shall propose a definition or algorithm for determining valid data. Such definition as agreed upon by the contractor and EPA shall assist in determining achievement of the contract objectives.

Figure 1.4.11.1 (continued). Quality assurance information requested in an RFP for a major air pollution monitoring network.

1. First article configuration inspection (FACI). Purchase contracts that involve multiple instrument systems of identical design and/or monitoring stations of identical design should require a FACI as part of the contract. For further details see Section 1.4.19 (Configuration Control).

2. Systems acceptance test (SAT). The purchase of elaborate monitoring systems that may involve pollutant and meteorological sensing instruments, telemetry and computer operations should include an extensive performance check-out before the system is accepted. Requirements for a SAT should be included in the contract. An example of a SAT and a FACI in contracts for major air pollution monitoring networks operated by the U.S. Environmental Protection Agency is described in Reference 1.

3. Certificate of analysis. The purchase of materials used in the measurement system calibration (e.g., pollutant cylinder gases) and materials used for sample collection or analysis whose specifications are critical, should include a certificate of analysis with the purchase. The certificate should include the certified result for characteristic measured, method of analysis, name of analyst and, where applicable, traceability of the measurement to a material or standard of higher accuracy than that purchased. See Section 1.4.17 for further details on traceability.

4. Reliability testing. See Section 1.4.20 for further details.

Procurement specifications and acceptance procedures for materials and equipment - The most common problems experienced in the procurement of air pollution measurement systems or parts thereof deal with inadequate and incorrect specifications for special purchases which may or may not be off-the-shelf items. Recommended acceptance limits for materials and equipment for each measurement method are included in Volume II -Ambient-Air Specific Methods and Volume III - Stationary-Source Specific Methods in the method activity matrices. These recommended acceptance limits should be reviewed, and where appropriate, procurement specifications should be written into the procurement contract.

The procurement document for materials and equipment should include the detailed specifications for the items and specific procedures for the measurement test to determine acceptability of the items.

In the quality acceptability provision of a procurement contract for materials, the following information should be included:

1. List of quality characteristics to be tested.
2. Method of measurement or test for each characteristic.
3. Specific procedure for sampling to obtain the materials or items to measure or test.
4. Specific acceptability limits.

An example of such specifications and acceptance procedures for glass-fiber filters used in the reference method for the determination of suspended particulates in the atmosphere (High Volume Method) is shown in Figure 1.4.11.2. These specifications and acceptance procedures were included in an U.S. Environmental Protection Agency contract.

Procurement specifications for warranty, inspection and corrective action for materials and equipment - The procurement contract should specify a warranty to protect the purchaser and where appropriate an inspection and corrective action procedure for procurement acceptance. By way of example, samples of optional clauses for procurement contracts of the U.S. Environmental Protection Agency are shown in Figure 1.4.11.3. Optional Clause 1 is a warranty for fixed price off-the-shelf procurement contract for materials and equipment. Commonly, procurement contracts are written for the design, development and delivery of a special air-pollution sensor or measurement system. If a cost-plus-fixed-fee contract is written, Optional Clause 2 may be included in the contract.

Procurement documentation by vendors - Purchase contracts should instruct the vendor to mark containers or reagents and/or packing slips with the following information when applicable:

1. Purchase order number.
2. Name of material.
3. Vendor's name and address.

The contractor shall deliver 300,000 high purity glass-filters made from a minimum number of batches of slurry. These filters shall measure 8" ± 1/16" by 10" ± 1/16". The filters shall be numbered front and back on the outer 1/2" of the filter. The filters shall be packaged in sets of 65 filters contained in non-contaminating plastic jackets which are sandwiched between cardboard sheets and inserted into envelopes suitable for mailing. These filters are to be used for particulate sampling nationwide and shall retain their integrity and tear strength up to 400 degrees C. The filter shall be made with the machine direction parallel to the 8" side of the sheet.

The filters shall not come apart when folded lengthwise one time; they shall also have a tensile strength (measured on 15-mm-wide specimens cut from at least 250 filters) of at least 500 grams/15 mm strip when pulled at a rate of 1/2 in./min with 50 mm of specimen between the pulling jaws. Tensile strength measurements shall be made in the machine direction.

The filter shall give at least 99.9% retention for particles of 0.3 µ at a flow of 60 cubic feet/minute as determined by a DOP test according to ASTM Method 2986. The filters shall not have any pinholes detectable visually on a viewing screen and shall weigh at least 3.5 grams but no more than 4.5 grams.

The contractor shall identify by filter numbers the production units used to make this run of filters.

The vendor shall establish his ability to supply the requisite filters in a reasonable time frame, not to exceed 120 days from date of award of contract.

The vendor shall furnish data on the filters to be supplied as shown in the table on acceptability limits. This data is to be gathered from an independent laboratory and 130 filters from this group being sent to EPA for corroborative analysis. The data as shown in the table will be supplied to EPA prior to final financial negotiation of the contract. The lot of 300,000 filters will be accepted by EPA as meeting specifications if the lot of samples drawn meet the acceptance testing set forth in the protocol for "Filter Acceptance."

Figure 1.4.11.2. Procurement contract for glass-fiber filters - specifications.

Elements of Ions	Max. conc. in g/8 x 10 filter	Method of determination
SO_4^{-2}	300	1
NO_3^{-1}	100	1
NH_4^+	50	2
F^-	150	3
Cl^-	2,000	1
Benzene organic extract	100	4
Antimony (Sb)	50	5,6
Arsenic (As)	70	7
Benzo(a)pyrene (BaP)	0.2	4
Beryllium (Be)	1	5,6
Bismuth (Bi)	35	5,6
Cadmium (Cd)	5	5,6
Chromium (Cr)	10	5,6
Cobalt (Co)	10	5,6
Copper (Cu)	7	5,6
Iron (Fe)	250	5,6
Lead (Pb)	30	5,6
Manganese (Mn)	10	5,6
Molybdenum (Mo)	15	5,6
Nickel (Ni)	10	5,6
Tin (Sn)	15	5,6
Titanium (Ti)	10	5,6
Vanadium (V)	20	5,6
Zinc (Zn)	90	5,6
pH	Below 9.6	8

Figure 1.4.11.2 (continued). Procurement contract for glass-fiber filters - acceptability limits.

The Environmental Protection Agency, using the 130 sample filters shall:

1. Weigh all filters - reject if over 13 are over 4.5 or below 3.5 grams.
2. Check all filters on a light-table. Reject if over 5 have pin holes.
3. Check all filters for size. Reject if more than 10 do not meet size specifications.
4. Check random 5 filters for tensile strength. Reject if any do not meet spcifications.
5. Check 5 filters for DOP. Reject if any do not meet specifications.
6. Subject the remaining 120 filters, that is, those not destructively tested in 4 and 5 above, to chemical analysis. Reject if over 5 do not meet specifications.

Figure 1.4.11.2 (continued). Procurement contract for glass-fiber filters - acceptability limits.

Optional Clause 1 - Warranty for fixed price off-the-shelf procurements

The Contractor agrees that the supplies or services furnished by the Contractor under this contract shall be covered by warranties which are either standard or customary in the trade or are substantially similar to, and not in excess of, a standard or customary trade warranty. The Contractor shall furnish any warranty description at time of delivery. Supply packages shall be simply marked to show existence of the warranty, its expiration date and the company official to be notified.

Optional Clause 2 - Inspection and correction of defects for cost-plus-fixed-fee procurements

(a) All work, under this contract shall be subject to inspection and test by the Government (to the extent practicable) at all times (including the period of performance) and places, and in any event prior to acceptance. The Contractor shall provide and maintain an inspection system acceptable to the Government covering the work, hereunder. The Government, through any authorized representative, may inspect the plant or plants of the Contractor or of any of his subcontractors engaged in the performance of this contract. If any inspection or test is made by the Government on the premises of the Contractor or a subcontractor the Contractor shall provide and shall require subcontractors to provide all reasonable facilities and assistance for the safety and convenience of the Government inspectors in the performance of their duties. All inspections and tests by the Government shall be performed in such a manner as will not unduly delay the work. Except as otherwise provided in this contract, final inspection, and acceptance shall be made at the place of delivery as promptly as practicable after delivery and shall be deemed to have been made no later than ninety (90) days after the date of such delivery, if acceptance has not been made earlier within such period.

Figure 1.4.11.3 EPA optional clauses for procurement of materials and equipment (2)

(b) At any time during performance of this contract, but not later than six (6) months (or such other time as may be provided in the contract) after acceptance of all of the end items (other than designs, drawings, or reports) to be delivered under this contract the Government may require the Contractor to remedy by correction or replacement, as directed by the Contracting Officer, any failure by the Contractor to comply with the requirements of this contract. Any time devoted to such correction or replacement shall not be included in the computation of the period of time specified in the preceding sentence, except as provided in (d) below. Except as otherwise provided in paragraph (c) below, the allowability of the cost of any such replacement or correction shall be determined as provided in the clause of this contract entitled "Allowable Cost, Fixed Fee, and Payment," but no additional fee shall be payable with respect thereto. Corrected articles shall not be tendered again for acceptance unless the former tender and the requirement of correction is disclosed. If the Contractor fails to proceed with reasonable promptness to perform such replacement or correction, the Government (i) may by contract or otherwise perform such replacement or correction and charge to the Contractor any increased cost occasioned the Government thereby, or may reduce any fixed fee payable under the contract (or require repayment of any fixed fee theretofore paid) in such amount as may be equitable under the circumstances, or (ii) in the case of articles not delivered, may require the delivery of such articles, and shall have the right to reduce any fixed fee payable under this contract (or to require repayment of any fixed fee theretofore paid) in such amount as may be equitable under the circumstances, or (iii) may terminate this contract for default. Failure to agree to the amount of any such increased cost to be charged to the Contractor or to such reduction in, or repayment of, the fixed fee shall be deemed to be a dispute concerning a question of fact within the meaning of the clause of this contract entitled "Disputes."

(c) Notwithstanding the provisions of paragraph (b) above, the Government may at any time require the Contractor to remedy by correction or replacement, without cost to the Government, any failure by the Contractor to comply with the requirements of this contract, if such failure is due to fraud, lack of good faith or willful misconduct

Figure 1.4.11.3 (continued). EPA optional clauses for procurement of materials and equipment (2).

on the part of any of the Contractor's directors or officers, or on the part of any of his managers, superintendents, or other equivalent representatives, who has supervision or direction of (i) all or substantially all of the Contractor's business, or (ii) all or substantially all of the Contractor's operations at any one plant or separate location in which this contract is being performed, or (iii) a separate and complete major industrial operation in connection with the performance of this contract. The Government may at any time also require the Contractor to remedy by correction or replacement, without cost to the Government, any such failure caused by one or more individual employees selected or retained by the Contractor after any such supervisory personnel has reasonable grounds to believe that any such employee is habitually careless or otherwise unqualified.

(d) The provisions of paragraph (b) above shall apply to any corrected or replacement end item or component until six months after its acceptance.

(e) The Contractor shall make his records of all inspection work available to the Government during the performance of this contract and for such longer period as may be specified in this contract.

(f) Except as provided in this clause and as may be provided in the contract, the Contractor shall have no obligation or liability to correct or replace articles which at the time of delivery are defective in material or workmanship or otherwise not in conformity with the requirements of this contract.

(g) Except as otherwise provided in the contract, the Contractor's obligation to correct or replace Government furnished property (which is property in the possession of or acquired directly by the Government and delivered or otherwise made available to the Contractor) shall be governed by the provisions of the clause of this contract entitled "Government Property."

Figure 1.4.11.3 (continued). EPA optional clauses for procurement of materials and equipment (2).

4. Vendor's lot number.
5. Quantity.
6. Material specification number and date.

This assures that the material is properly identified and that the vendor is using the latest specifications.

Receiving inspections - Procurement specifications should be checked upon receipt of purchases. The method activity matrices in the specific methods in Volume II and III show the following in regard to procurement specifications:

1. Acceptance criteria
2. Frequency of check or measurement.
3. Method of measurement.

The frequency of checking incoming materials can be determined and controlled by statistical applications.

Quality system survey - These surveys should be made prior to purchase of equipment or material involving large outlays of capital or the award of a large service contract. The purpose is to determine the quality assurance system planned by the vendor or contract for the procurement. In the case of items that are purchased in large quantities over a period of time, repeat visits may be made to ensure continuing production of material of satisfactory quality.

Vendor rating system - When material has been purchased over a period of time, the information developed on the vendor becomes a history of the quality performance of that vendor. This information can be used to rate the vendor against an established performance standard or against competitors. A simple, yet effective vendor rating system is described in Reference 3.

Not only is it important to identify the quality vendors (by means of a rating system) but it is equally important to identify poor quality vendors.

1.4.11.3 REFERENCES

1. Von Lehmden, D. J., R. C. Rhodes, and S. Hochheiser. "Applications of Quality Assurance in Major Air Pollution Monitoring Studies - CHAMP and RAMS," for publication in the proceedings of the International Conference on Environmental Sensing and Assessment, Las Vegas, Nevada, September 1975. Copies available from Quality Assurance Branch, EPA, Research Triangle Park, North Carolina 27711.

2. EPA, Office of Administration, Contracts Management Division, Research Triangle Park, North Carolina 27711.

3. Siegel, J. A Practical Approach to Vendor Ratings, Industrial Quality Control 23, July 1966. pp. 17-19.

BIBLIOGRAPHY

1. Covino, C.P. and Meghri, A.W. Quality Assurance Manual. Industrial Press, Inc., New York. pp. 36-37b. 1967.

2. Juran, J.M., (ed.). Quality Control Handbook. 2nd Ed. McGraw-Hill, New York. 1962. Section 24, pp. 2-37.

3. Juran, J.M. and Gryna, F.M. Quality Planning and Analysis. McGraw-Hill, New York. 1970. Chapters 19 and 20, pp. 402-433.

4. Feigenbaum, A.V. Total Quality Control. McGraw-Hill, New York. 1961. Chapter 16, pp. 481-520.

1.4.12 CALIBRATION

1.4.12.1 ABSTRACT

1. A calibration plan should be developed and implemented for measuring and test equipment and calibration standards to include: calibration intervals; listing of required calibration standards, including recommended mechanism by which to establish traceability to standards available at the National Bureau of Standards or other fundamental standards; required environmental conditions for calibration; and a calibration record system.

2. Calibration standards quality should have 4 to 10 times the accuracy of measuring and test equipment being calibrated. A listing of environmental standards available from the National Bureau of Standards is shown. Calibration gases purchased from commercial vendors need traceability to available certified gases from the National Bureau of Standards.

3. Adequate environmental conditions must be provided during calibration. Recommended environmental conditions for temperature, relative humidity, electric power, etc. are given.

4. A calibration interval for recalibration should be assigned to all measuring and test equipment and calibration standards. A record of calibrations should be maintained.

5. Written calibration procedures should be provided for all measuring and test equipment and calibration standards. A procedure should:

a. Specify where the procedure is applicable.

b. Provide a brief description of the calibration principle.

c. List calibration standards and accessory equipment required.

d. Specify the record system including example data sheets.

6. Calibration sources performing calibrations should prepare reports that include showing the traceability of the calibration standard used in the calibration.

7. Measuring and test equipment should be labeled to indicate calibration date and when the calibration expires.

8. A calibration checklist is provided that may be used as a guide on calibration during laboratory evaluations.

1.4.12.2 DISCUSSION

General - Calibration procedures require the application of primary or secondary standards. The standards used, whether they are apparatus or reagent standards, should be certified as being traceable to standards of the National Bureau of Standards, or some other recognized fundamental standard. This kind of traceability is possible even when standards are generated in the laboratory. Regardless of the type of calibration equipment or material, an effective quality assurance program requires accuracy levels of these materials that are consistent with the method of analysis.

The calibration policies and procedures set forth within this section should apply to all measuring and test equipment identified, including:

1. Sampling equipment at sampling stations.
2. Analytical equipment in the laboratory.
3. Flow rate (e.g., rotameters), volume (e.g., dry gas meters), pressure, vacuum and temperature measurement equipment both at the sampling stations and in the laboratory.

Calibration Plan - A detailed plan should be provided for controlling the accuracy of measuring and test equipment and calibration standards utilized. The plan should include:

1. Realistic calibration intervals for measuring and test equipment and, for each calibration standard, designation of satisfactory calibration sources.
2. A listing of all required calibration standards with proper nomenclature and identification numbers assigned.
3. The mechanism used to establish traceability of calibration standards to standards available at the National Bureau of Standards or other recognized fundamental standards.

4. The environmental conditions (temperature, relative humidity, pressure, etc.) to be provided by the calibration source under which the calibration standards will be utilized and the calibration performed.

5. Written calibration procedures for measuring and test equipment and calibration standards, including document control numbers for reference purposes.

6. A description of a calibration record system, including samples of labels, decals, record cards, etc.

Calibration Standards Quality[1] - Transfer standards should have 4 to 10 times the accuracy of field and laboratory instruments and gauges. For example, if a thermometer used in the field to determine air temperature has a specified accuracy of ± 2°F, it should be calibrated against a laboratory thermometer with an accuracy of at least ± 0.5°F. The calibration standards used in the measurement system should be calibrated against higher-level, primary standards having unquestionable and higher accuracy. The higher-level standards should be certified by the National Bureau of Standards or another recognized organization or derived from accepted values of physical or chemical constants. A listing of environmental standards available from the National Bureau of Standards is shown in Figure 1.4.12.1. These environmental standards may be purchased from the National Bureau of Standards, Office of Standard Reference Materials, Washington, D.C. 20234.

SRM No.	Type	Unit
1604a	Oxygen in N_2, 1.5 ppm	cyl
1607	Oxygen in N_2, 212 ppm	cyl
1609	Oxygen in N_2, 20.95 Mol. %	cyl
1621	Sulfur in residual fuel oil, 1.05 wt %	100 ml
1622	Sulfur in residual fuel oil, 2.14 wt %	100 ml
1623	Sulfur in residual fuel oil, 0.268 wt %	100 ml
1624	Sulfur in distillate fuel oil, 0.211 wt %	100 ml
1625	Sulfur dioxide permeation tube, 10 cm, 2.8 μg/min at 25°C	ea
1626	Sulfur dioxide permeation tube, 5 cm, 1.4 μg/min at 25°C	ea
1627	Sulfur dioxide permeation tube, 2 cm, 0.56 μg/min at 25°C	ea
1629	Nitrogen dioxide permeation device, individual between 0.5 to 1.5 μg/min at 25°C	ea
1631	Sulfur in coal, three concentrations	set (15)
1632	Trace elements in coal	75 g
1633	Trace elements in coal fly ash	75 g
1634	Trace elements in fuel oil	IN PREPARATION
1636	Lead in Reference Fuel, four concentrations, 0.03, 0.05, 0.07 and 2.0 g/gal	set (12)
1637	Lead in Reference Fuel, three concentrations, 0.03, 0.05 and 0.07 g/gal	set (12)
1638	Lead in Reference Fuel, one concentration, 2.0 g/gal	set (12)
1665	Propane in air, 2.8 ppm	cyl
1666	Propane in air, 9.5 ppm	cyl
1667	Propane in air, 48 ppm	cyl
1668	Propane in air, 95 ppm	cyl
1669	Propane in air, 475 ppm	cyl
1673	Carbon dioxide in N_2, 0.95 Mol. %	cyl
1674	Carbon dioxide in N_2, 7.2 Mol. %	cyl
1675	Carbon dioxide in N_2, 14.2 Mol. %	cyl
1677	Carbon monoxide in N_2, 9.74 ppm	cyl
1678	Carbon monoxide in N_2, 47.1 ppm	cyl
1679	Carbon monoxide in N_2, 94.7 ppm	cyl
1680	Carbon monoxide in N_2, 484 ppm	cyl
1681	Carbon monoxide in N_2, 957 ppm	cyl
1683	Nitric oxide in N_2, 50 ppm	cyl
1684	Nitric oxide in N_2, 100 ppm	cyl
1685	Nitric oxide in N_2, 250 ppm	cyl
1686	Nitric oxide in N_2, 500 ppm	cyl
1687	Nitric oxide in N_2, 1000 ppm	cyl

Figure 1.4.12.1 Environmental Standard Reference Materials available from the National Bureau of Standards in 1975

Calibration gases purchased from commercial vendors normally contain a certificate of analysis. Whenever a certified gas is available from the National Bureau of Standards, commercial gas vendors should be requested to establish traceability of the certificate of analysis to this certified gas. Inaccurate concentrations on certificates of analysis may result in serious errors in reported measured pollutant concentrations. By way of example, an EPA report of the collaborative test results for carbon monoxide using reference method 10 (test method for Federal regulations on Standards of Performance for New Stationary Sources) includes the following statement in the Summary:

> "Some calibration gas suppliers provide certificates of analysis that show errors of as much as 30% when compared with standard gases."[2]

Environmental Conditions - Measuring and test equipment and calibration standards should be calibrated in an area that provides control on environmental conditions to the degree necessary to assure required accuracy. The calibration area should be reasonably free of dust, vapor, vibration, and radio frequency interferences; and it should not be close to equipment that produces noise, vibration, or chemical emissions or close to areas in which there is chemical testing or use of high-power radar work, etc.

The laboratory calibration area should have adequate temperature and humidity controls. A temperature of 68 to 73°F and a relative humidity of 35 to 50 percent normally provide a suitable environment.

A filtered air supply is desirable in the calibration area. Dust particles are more than just a nuisance; they can be abrasive, conductive, and damaging to instruments.

Other environmental conditions for consideration are:

1. Electric power. Recommended requirements for electrical power within the laboratory should include voltage regulation of at least 10 percent (preferably 5 percent); low values of harmonic distortion; minimum voltage fluctuations caused by interaction of other users on main line to laboratory(separate input power if possible); and a suitable grounding system established to assure equal potentials to ground throughout the laboratory.

2. Lighting. Adequate lighting (suggested values--80- to 100-foot candles) should be provided for workbench areas. The lighting may be provided by overhead incandescent or fluorescent lights. Fluorescent lights should be shielded properly to reduce electrical noise.

Calibration Intervals[1] - All calibration standards and measuring and test equipment should be assigned an established interval for calibration (Figure 1.4.12.2). In the absence of an established calibration interval (based on equipment manufacturer's recommendation, authorized Government specifications, etc.) for a particular item, an initial

CALIBRATION FREQUENCY SCHEDULE

Type	Minimum Frequency	Instruction Number or Calibration Method

Figure 1.4.12.2 Calibration frequency schedule

servicing interval should be assigned by the laboratory or calibrating service. The servicings should be specified in terms of time or, in the case of certain types of test and measuring equipment, a usage period.

The establishment of significant intervals should be based on inherent stability, purpose or use, accuracy, and degree of usage. The intervals may be shortened or lengthened by evaluating the results of the previous and present calibrations and adjusting the schedule to reflect the findings. These evaluations must provide positive assurance that calibration interval adjustments will not adversely affect the accuracy of the system.

The laboratory should maintain proper usage data and historical records for all equipment, to ascertain whether an adjustment of the calibration interval is warranted.

Adherence to the calibration frequency schedule is mandatory. It is operated as follows: as the interval is established, the Calibration Control Card (Figure 1.4.12.3) is prepared and flagged, at the appropriate place on the calendar printed at the top of the card, using 12 different color flags for months or using two flags for month and date. These cards are sorted and filed by sequence of date, so that the cards that mature on a given day identify which instruments are to be recalled for check. Notice of recall is given prior to the due date. It may be necessary to

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.

INSTRUMENT RECORD CARD

Instr. type__________ Ident. No.__________ Instr. Code No.__________

Model No.__________ Ser. No.__________

Date rec'd.__________ Date insp.__________ O.K.'d by__________

Checking interval__________ App'd. by__________

Location__________ Rec'd. by__________

Calibration Responsibility__________

Calibration Instruction Number

FRONT

Figure 1.4.12.3 Calibration control card

Date checked	Checked By	Results of check	Date checked	Checked By	Results of check

REVERSE

Figure 1.4.12.3 (Continued) Calibration control card calibrate between normal calibration dates if there is evidence of inaccuracy or damage.

Calibration Procedures - Written step-by-step procedures for calibration of measuring and test equipment and use of calibration standards should be provided by the laboratory in order to eliminate possible measurement inaccuracies due to differences in techniques, environmental conditions, choice of higher-level standards, etc. These calibration procedures may be prepared by the laboratory, or the laboratory may use published standard practices or written instructions that accompany purchased equipment. These procedures should include the following information:

1. The specific equipment or group of equipment to which the procedure is applicable. ("Like" equipment or equipment of the same type, having compatible calibration points, environmental conditions, and accuracy requirements, may be serviced by the same calibration procedure.)

2. A brief description of the scope, principle, and/or theory of the calibration method.

3. Fundamental calibration specifications, such as calibration points, environmental requirements, and accuracy requirements.

4. A list of calibration standards and accessory equipment required to perform an effective calibration. Manufacturer's name, model number, and accuracy should be included as applicable.

5. A complete procedure for calibration arranged in a step-by-step manner, clearly and concisely written.

6. Calibration procedures should provide specific instructions for obtaining and recording the test data, and should include data sheets.

When available, published procedures may be used. NBS Handbook 77, Precision Measurement and Calibration,[3] published by the National Bureau of Standards, provides calibration procedures for many types of electrical, hydraulic, electronic, and mechanical measuring instruments and can be used when available.

Many calibration procedures require statistical analysis of results. A detailed sample of computations for calibration of an NO_2 monitor is provided in Appendix J.

Calibration Source[1] - All calibrations performed by or for the laboratory should be traced through an unbroken chain (supported by reports or data sheets) to some ultimate or national reference standards maintained by a national organization such as the National Bureau of Standards. The ultimate reference standard can also be an independent reproducible standard, i.e., a standard that depends on accepted values of natural physical constants.

An up-to-date report for each calibration standard used in the calibration system should be provided. If calibration services are performed by a commercial laboratory on a contract basis, copies of reports issued by them should be maintained on file.

All reports should be kept in a suitable file and should contain the following information:

1. Report number.

2. Identification or serial number of the calibration standard to which the report pertains.

3. Conditions under which the calibration was performed (temperature, relative humidity, etc.).

4. Accuracy of calibration standard (expressed in percentage or other suitable terms).

5. Deviation or corrections.

6. Corrections that must be applied if standard conditions of temperature, etc., are not met or differ from those at place of calibration.

Contracts for calibration services should require the commercial laboratory to supply records on traceability of their calibration standards.

Labeling[1] - All equipment to be calibrated should have affixed to it in plain sight a tag bearing the following information:

Description: ____________________

Ident. No. ______________________

Last Calibrated: _________________

Calibrated By: ___________________

Calibration Expires: ______________

NOTE: Use of this instrument beyond the calibration expiration date is prohibited.

When the equipment size or its intended use limits the application of labels, an identifying code should be applied.

Equipment past due for calibration should be removed from service either physically or, if this is impractical, should be impounded by tagging or other means.

Surveying[4] - The Calibration Checklist (Figure 1.4.12.4) is intended for use as a guide in evaluation of laboratories on calibration procedures for measuring and test equipment.

Omission of one or several of these check-points does not necessarily imply that a given laboratory does not maintain required accuracy in its calibration program; rather, the overall system should be appraised in light of its ability to provide such accuracy.

DATE:____________________

Location:

1. Is each piece of measuring or test equipment calibrated on a routine, periodic, basis against appropriate standards traceable to the National Bureau of Standards?

Yes_____No _____Remarks:____

2. Are schedules for calibration established for each type of instrumentation equipment?

Yes_____No______Remarks:____

Have the following factors been considered in designing the schedule:

a. Severity of environment. Yes_____No_____
b. Severity of use. Yes_____No_____
c. Frequency of use. Yes_____No_____
d. Delicacy of the measuring equipment. Yes____No_____
e. Accuracy of measurement required. Yes____No_____
f. Calibration history. Yes_____No_____
g. Importance of the characteristic or condition measured. Yes____No_____
h. Sensitivity of the test methods used on the type of equipment.
Yes____No____Remarks:____________________________________

3. Is calibration performed to: published standard practices, manufacturers' written instructions, or other approved instructions?

Yes______No______Remarks:____

4. Is there positive identification of calibration status of each unit of equipment provided by color code, tag, label, stamp, etc.?

Yes_____No______ Remarks:____

5. Does the calibration status identification include the following?

a. Date of calibration. Yes_____No_____
b. Identification of technician performing calibration. Yes____No____
c. Due date for expiration of calibration status. Yes____No____

Remarks:____

6. Is calibration conducted in an environment controlled to the extent necessary to assure continuation of measurements of the required accuracy?

Yes______No_____Remarks:____

Figure 1.4.12.4 Calibration checklist

7. Does environmental control include controls on:
   a. Temperature Yes______No______
   b. Humidity Yes______No______
   c. Vibration Yes______No______
   d. Dust and/or gaseous contaminants. Yes______No______
   e. Cleanliness. Yes______No______

   Remarks:____

8. Are compensation corrective factors applied to calibration results obtained in an environment department from standard conditions?

Yes______No______Remarks:______________________________

9. Are there individual instrument calibration record cards: Yes______No______
Do they show:

   a. Description Yes______No______
   b. Identification number Yes______No______
   c. Calibration interval Yes______No______
   d. Reference to calibration instruction Yes______No______
   e. Identification of calibrating technician Yes______No______

   Remarks:____

   f. Date of calibration Yes______No______
   g. Calibration expiration due date Yes______No______
   h. Actual values obtained Yes______No______
   I. Location of equipment Yes______No______

   Remarks: ____

____

   Recommendations:____

Signed________________________________

____
Organization

Figure 1.4.12.4 (Continued) Calibration checklist

1.4.12.3 REFERENCES

1. Evaluation of Contractor's Calibration System. " MIL-HDBK-52, Department of Defense, Washington, D.C. July 1964.

2. Collaborative Study of Method 10-Reference Method for Determination of Carbon Monoxide Emissions from Stationary Sources. EPA, National Environmental Research Center, Research Triangle Park, N. C. EPA report 650/4-75-001. January 1975. 80 pp.

3. Precision Measurement and Calibration. National Bureau of Standards, Washington, D.C. NBS Handbook 77.

4. Industrial Hygiene Service Laboratory Quality Control Manual. National Institute for Occupational Safety and Health. Cincinnati, Ohio. Technical Report No. 78. 1974.

BIBLIOGRAPHY

1. Beckwith, T.G., and Buck, N.L. *Mechanical Measurements*. Addison-Wesly Publishing Company, Reading, Massachusetts. 1969. Chapter 10.

2. Covino, C.P., and Meghri, A.W. *Quality Assurance Manual*. Industrial Press, Inc., New York. 1967.

3. Feigenbaum, A.V. *Total Quality Control*. McGraw-Hill Book Company, New York. 1961. pp. 512-514.

4. Kennedy, C.W., and Andrews, D.E. *Inspection and Gauging*. 4th ed. Industrial Press, Inc., New York. 1967. Chapter 14.

5. Calibration System Requirements. MIL-C-45662A, Department of Defense, Washington, D.C. February 1962.

1.4.13 CORRECTIVE ACTION

1.4.13.1 ABSTRACT

1. Corrective actions are of two kinds:

    a. Corrective action - on-the-spot or immediate, to correct or repair nonconforming data or equipment.

    b. Corrective action - long-term, to eliminate causes of nonconformance.

2. Steps comprising a closed-loop corrective action system are:

    a. Define the problem.

    b. Assign responsibility for investigating the problem.

    c. Investigate and determine the cause of the problem.

    d. Determine a corrective action to eliminate the problem.

    e. Assign and accept responsibility for implementing the corrective action.

    f. Establish effectiveness of the corrective action and implement the correction.

    g. Verify that the corrective action has eliminated the problem.

Corrective action procedures recognize the need for an assigned individual to test continually the effectiveness of the system and the corrective actions.

1.4.13.2 DISCUSSION

On-the-spot or immediate corrective action - This is the process of correcting malfunctioning equipment.

In a quality assurance program, one of the most effective means of preventing trouble is to respond immediately to reports from the operator of suspicious data or equipment malfunctions. Application of proper corrective actions at this point can reduce or prevent the collection of poor-quality data. Established procedures for corrective actions are available in the methods if the performance limits are found to be exceeded (either through direct observation of the parameter or through review of control charts). Specific control procedures, calibration, pre-sampling or pre-analysis operational checks, etc., are designed to detect instances in which corrective action is necessary. A checklist for logical alternatives for tracing the source of a sampling or analytical error is provided to the operator. Trouble-shooting guides for operators (field technicians or lab analysts) are generally found in instrument manufacturer's manuals. On-the-spot corrective actions routinely made by field technicians or lab analysts should be documented as normal operating procedures, and no specific documentation other than notations in operations logbooks need be made.

Long-term corrective action - The purpose of long-term corrective action is to identify and eliminate causes of nonconformance; hopefully they will be eliminated permanently. To improve data quality to an acceptable level and to maintain data quality at an acceptable level, it is necessary that the quality assurance system be sensitive and timely in detecting out-of-control or unsatisfactory conditions. It is equally important that, once the conditions of unacceptable quality data are indicated, a systematic and timely mechanism is established to assure that the condition is reported to those who can correct it and that a positive loop mechanism is established to assure that appropriate corrective action has been taken. For major problems it is desirable that a formal system of reporting and recording of corrective actions be established.

Closed-loop corrective action system for major problems - Experience has shown that most problems will not disappear until positive action has been taken by management. The significant characteristic of any good management system is the step that closes the loop--the determination to make a change if the system demands it.

The following discussion outlines the considerations and procedures necessary to understand and implement an effective closed-loop corrective action system for major problems. Effective corrective action occurs when many individuals and departments cooperate in a well planned program. There are several essential steps that must be taken to plan and implement a corrective action program that achieves significant results.

Corrective actions should be a continual part of the laboratory system for quality, and they should be formally documented. Corrective action is not complete until it is demonstrated that the action has effectively and permanently corrected the problem. Diligent follow-up is probably the most important requirement of a successful corrective action system.

Figure 1.4.13.1 illustrates the sequence of activities involved in operating a closed-loop corrective action system.

Initiation, use, and completion of the corrective action request - The corrective action request (Figure 1.4.13.2) may be initiated by any individual who observes a major problem. The corrective action request should be limited to a single problem. If more than one problem is involved, each should be documented on a separate corrective action request form.

Use of a Master Log - Corrective action can be casual when the organization is small or the problems few. When this is not the case and the problems are severe and numerous, action documentation and status records are required. The primary tool of the corrective action system is the Corrective Action Request Form; the heart of the system is the Master Log (Figure 1.4.13.3). All problems to be in-

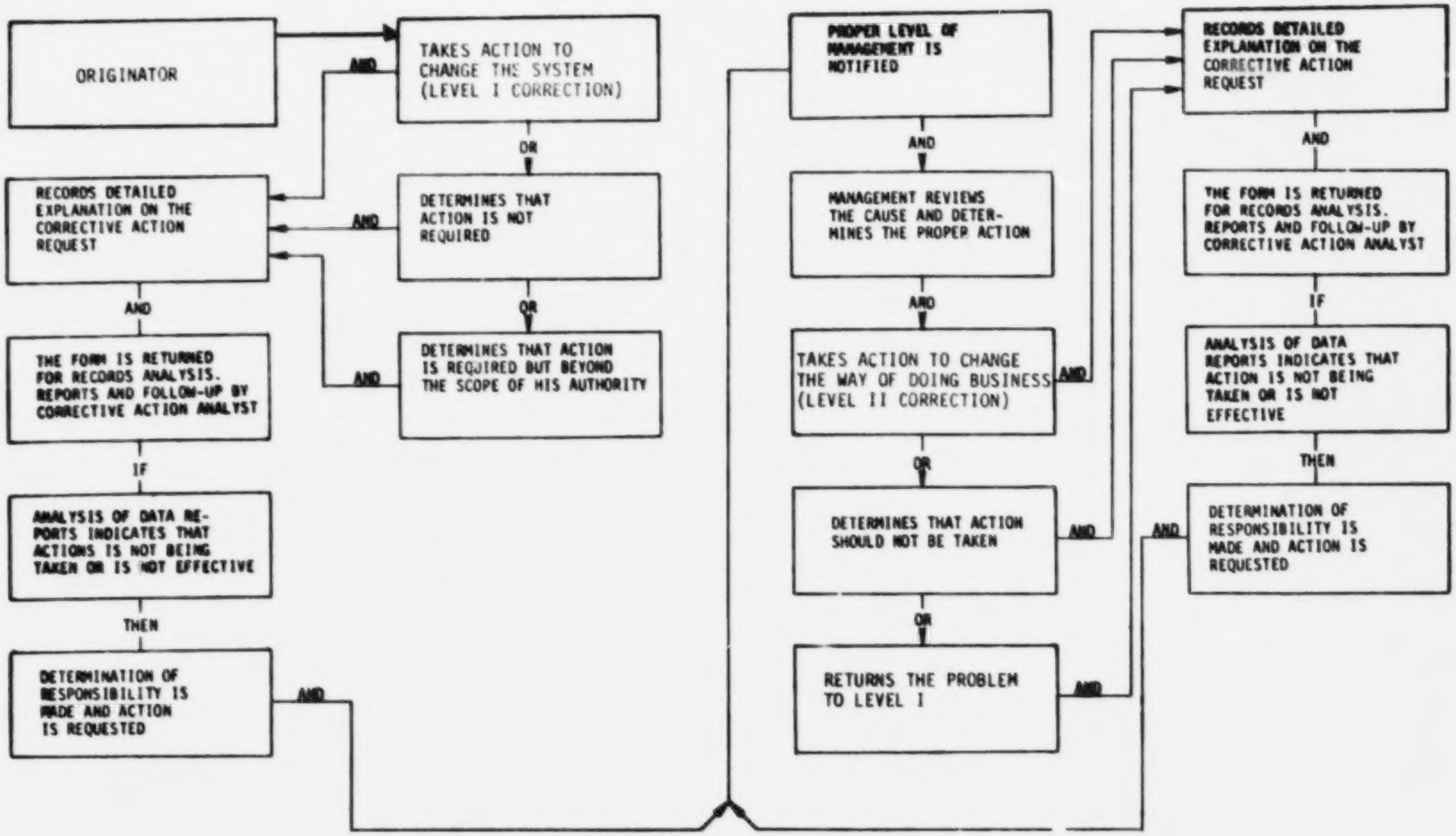

ORIGINATOR
AND
TAKES ACTION TO CHANGE THE SYSTEM (LEVEL I CORRECTION)
OR
DETERMINES THAT ACTION IS NOT REQUIRED
OR
DETERMINES THAT ACTION IS REQUIRED BUT BEYOND THE SCOPE OF HIS AUTHORITY
AND
AND
RECORDS DETAILED EXPLANATION ON THE CORRECTIVE ACTION REQUEST
AND
THE FORM IS RETURNED FOR RECORDS ANALYSIS. REPORTS AND FOLLOW-UP BY CORRECTIVE ACTION ANALYST
IF
ANALYSIS OF DATA REPORTS INDICATES THAT ACTIONS IS NOT BEING TAKEN OR IS NOT EFFECTIVE
THEN
DETERMINATION OF RESPONSIBILITY IS MADE AND ACTION IS REQUESTED
AND
PROPER LEVEL OF MANAGEMENT IS NOTIFIED
AND
MANAGEMENT REVIEWS THE CAUSE AND DETERMINES THE PROPER ACTION
AND
TAKES ACTION TO CHANGE THE WAY OF DOING BUSINESS (LEVEL II CORRECTION)
AND
OR
DETERMINES THAT ACTION SHOULD NOT BE TAKEN
AND
OR
RETURNS THE PROBLEM TO LEVEL I
AND
RECORDS DETAILED EXPLANATION ON THE CORRECTIVE ACTION REQUEST
AND
THE FORM IS RETURNED FOR RECORDS ANALYSIS. REPORTS AND FOLLOW-UP BY CORRECTIVE ACTION ANALYST
IF
ANALYSIS OF DATA REPORTS INDICATES THAT ACTION IS NOT BEING TAKEN OR IS NOT EFFECTIVE
THEN
AND
DETERMINATION OF RESPONSIBILITY IS MADE AND ACTION IS REQUESTED

Figure 1.4.13.1 A closed-loop corrective action system

Corrective Action Request Form No. ____________

NOTE: This form is part of a closed-loop corrective action system, and it is the originating organization's responsibility to follow the request through to resolution. Responsibility for a reply should be assigned to an individual rather than to an organization.

Originator ______________________________ Date __________________

Person responsible
for replying ______________________ Organization __________________

State nature of problem below:

PROBLEM IDENTIFICATION

____

____

____

____

____

State cause of problem: ____________________________________

____

____

____

____

Signature ______________________________

During what phase of operation was the problem identified? ________

RETURN

FROM: Name ____

Address ____

By: (Date) ______________________________________

Information copies to: (Names and addresses)

____

____

____

____

Figure 1.4.13.2 Corrective Action Request Form

Corrective Action Request Form No. __________

NOTE: Please return this form to the Corrective Action Analyst with the corrective action planned on or before the commitment date. If you determine that the action required is beyond the scope of your position, please indicate below who is responsible.

CORRECTIVE ACTION

Corrective Action Planned. Please indicate effectivity point and dates.

____

____

____

____

Is this a temporary action? Yes____ No____

Is this a permanent action? Yes____ No____

Signature________________________________ Date________

FINAL DISTRIBUTION

Original: CA Analyst

2nd Copy: Originator

3rd Copy: Return to Originator

4th, 5th, and 6th Copies Information

ACCEPTANCE OF CA

Initially Approved By:________

Date:______________

Suspense Data for Follow-up________

Final Approval By:________

Date:______________

THIS SPACE RESERVED FOR FOLLOW-UP

Figure 1.4.13.2 (Continued) Corrective Action Request Form

Corrective Action Request Form No.					
Date Submitted					
Date Received					
Submitted By					
General Problem Description					
Suspected Cause					
Assigned To					
Answer Due					
Nature of Problem Investigation					
Cause of Problem					
Corrective Action					
Corrective Action Responsibility					
Originator Notified					
Demonstration of Effectiveness					
Final Close-Out					

Figure 1.4.13.3 Corrective Action Master Log

cluded should be assigned a sequential number. This is accomplished by sequentially numbering the Corrective Action Request Form. If the problem is reported by letter, by memo, or by other means, a Corrective Action Request Form should be completed. All input documents should be entered into the Master Log for control purposes and for visibility to management.

1.4.14 QUALITY COSTS

1.4.14.1 ABSTRACT

Cost categories can be identified for a quality assurance system. By assigning costs according to quality assurance elements and grouping these by cost categories, cost effectiveness can be appraised.

1. Identification of costs is a prerequisite to cost reduction.

2. The American Society for Quality Control categorizes costs as: (a) prevention costs, (b) appraisal costs, (c) internal-failure costs, and (d) external-failure costs. For air pollution measurement systems, a more practical cost categorization is: (a) prevention cost, (b) appraisal costs, and (c) correction costs. The quality assurance elements listed in this Handbook have been placed in these three cost categories. Since accounting systems are not set up to accommodate cost breakdown by quality assurance elements, judgment is required to apportion the costs into the correct element and cost category.

3. Quality control cost figures should be reported quarterly to management.

4. Allocation of cost figures from the accounting system into the applicable cost categories helps to identify quality assurance elements whose costs may be disproportionate to benefits derived. Furthermore, quality cost figures provide input for budget forecasting.

1.4.14.2 DISCUSSION

Reasons for identifying quality assurance costs - Quality assurance costs should be identified and recorded primarily to identify elements of quality assurance whose costs may be disproportionate to the benefits derived. An additional purpose is to detect cost trends for budget forecasting.

Definition of quality cost categories[1,2] - The American Society for Quality Control lists four categories for quality costs:

1. Prevention costs associated with personnel engaged in designing, implementing, and maintaining the quality system. (Maintaining the quality system includes auditing the system.)

2. Appraisal costs associated with measuring, evaluating, or auditing of products, components, and purchased materials to assure conformance with quality standards and performance requirements.

3. Internal-failure costs associated with the manufacture of defective products, components, and materials that fail to meet quality requirements and thus result in manufacturing losses.

4. External-failure costs generated as a result of defective products having been shipped to customers.

However, since the quality assurance activities in environmental measurements (field sampling, laboratory analysis, data handling, etc.) will be different from those in manufacturing, it is more practical to categorize cost areas as follows:

1. Prevention costs spent keeping unacceptable data from being generated in the first place. Included here are such things as quality planning, training, etc.

2. Appraisal costs spent on efforts to maintain measurement system performance quality levels by formal audits of performance and interlaboratory and intralaboratory testing programs, etc.

3. Correction costs caused by occurrence of determinations or results that do not meet acceptance standards. These include voided data, spoiled filters, and duplicated tests.

The term "correction" is preferred rather than "failure" because of the close association the latter has to reliability.[3]

Elements of laboratory quality cost category - The elements of quality assurance discussed in this Handbook should be allocated to the three cost categories as follows:

1. Prevention costs
    a. Quality planning.
    b. Document control and revision (including measurement method write-ups).
    c. Training
    d. Configuration control.

  - e. Quality assurance plans for projects and programs.
  - f. Quality assurance manual.
  - g. Preventive maintenance.

2. Appraisal costs
  - a. Quality assurance associated with pretest preparation, sample collection, sample analysis, and data reporting.
  - b. Data validation.
  - c. Statistical analysis of data.
  - d. Procurement quality control.
  - e. Calibration.
  - f. Interlaboratory and intralaboratory testing.
  - g. Audit procedures.
  - h. Reliability.
  - i. Quality reports to management (including quality cost reports).

3. Correction costs

Corrective action (costs resulting from: additional sample collection and additional analysis because of invalid samples; reanalysis of collected samples because analysis was not controlled; scrapping of data because data are determined to be invalid; investigation of data quality problems; etc.)

As defined here, correction costs are those corrective action costs incurred by the organization generating the unacceptable data. Cost to others resulting from the use of this unacceptable data is not included as correction costs because this would be very difficult to quantitate.

Distribution of quality costs - After quality cost elements have been listed, it is necessary to allocate to them cost figures available from the accounting system. If actual costing figures are not available, then the accounting component should provide estimates that are as nearly accurate as possible.

It will usually be found that the costs are not uniformly distributed over the range of elements, a disproportionate percentage appearing in Appraisal and Correction costs. It has been found[4] that a relatively small increase in Prevention expenditures will yield large reductions in Appraisal and Correction costs. Therefore, the management should consider adoption of preventive measures to decrease total quality control costs.

Reporting quality costs - Quality control cost figures should be reported quarterly to management,[5] showing costs allocated to each of the three quality cost categories and relationships to total quality costs. It may also be useful to show quality costs by quality assurance element and/or against total organization expenditures.

Such reports should be presented as concisely as possible. Graphic presentations should be made where appropriate. An example of a quality cost graph is shown in Figure 1.4.14.1.

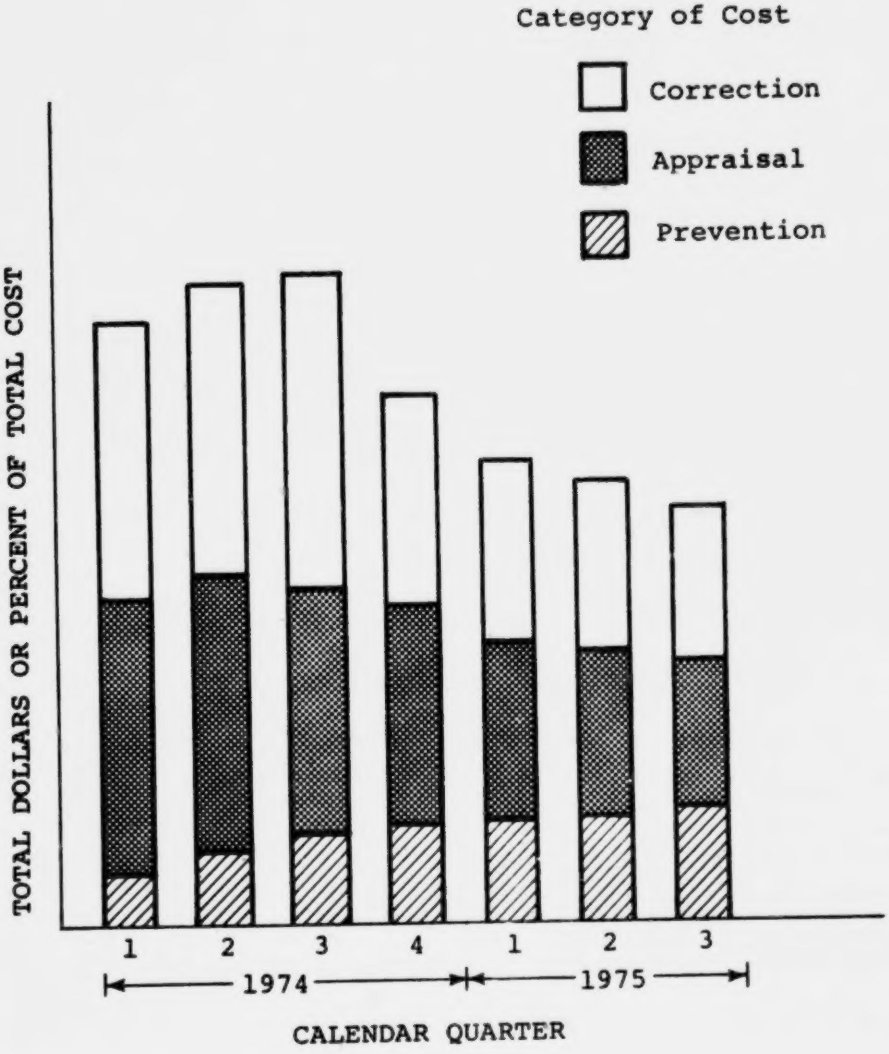


Figure 1.4.14.1 Quality cost by category.

This type of graph shows:

1. Expenditure by cost category.

2. Total quality cost expenditure. (One item to emphasize is the result increased prevention cost has on the total quality cost expenditure.)

3. Quarterly quality cost trend.

1.4.14.3 REFERENCES

1. Juran, J.M., and Gryna, F.M. Quality Planning and Analysis. McGraw-Hill, New York. 1970. Chapter 5, pp. 54-67.

2. "Quality Costs - What and How." American Society for Quality Control, Quality Cost Technical Committee. Milwaukee, Wisconsin. May 31, 1967.

3. Rhodes, R.C. "Implementing a Quality Cost System," Quality Progress. 5(2):16-19, February 1972.

4. Feigenbaum, A.V. Total Quality Control. McGraw-Hill, New York. 1961. Chapter 5, pp. 83-106.

5. Masser, W.J. "The Quality Manager and Quality Costs," Industrial Quality Control. XIV(4):5-8, October 1957.

1.4.15 INTERLABORATORY AND INTRALABORATORY TESTING

1.4.15.1 ABSTRACT

1. The purpose of interlaboratory testing is to identify those laboratories that are biased and to estimate variability (reproducibility) between laboratories.

2. Factors affecting interlaboratory testing include:

   a. Unbiased treatment of reference samples by the participating operator or analyst if it is apparent that the sample is a reference sample and not a routine sample.

   b. Stability of the reference samples during storage and shipping.

   c. Proper interpretation of the interlaboratory test procedure by the participating laboratories.

3. A special form of interlaboratory testing is collaborative testing. Collaborative study reports on ambient air and stationary-source measurement methods that are available from EPA are listed in Figure 1.4.15.1.

4. The purpose of intralaboratory testing is to identify sources of measurement error and to estimate their bias and variability (repeatability and replicability).

1.4.15.2 DISCUSSION

Interlaboratory testing - The purpose of interlaboratory testing is to identify those laboratories (and/or analysts) that are biased and to estimate measurement method reproducibility between laboratories. Laboratories participating in such programs are normally provided with standard reference samples and instructions for analysis. A coordinating laboratory prepares the reference samples and evaluates the results.

Interlaboratory test programs are highly recommended for ambient air and source emission measurement projects. These programs are particularly useful for the following type of projects:

1. Joint project - Whenever two or more laboratories work on a common project and each laboratory performs separate and independent measurements, an interlaboratory test program should be scheduled and repeated as often as necessary. An example of such a project is the Regional Air Pollution Study (RAPS) in the greater St. Louis area, in which several government agencies, universities, and private organizations are participating in a multiyear study of the atmosphere.[1]

2. Independent projects with similar goals - Whenever two or more laboratories working independently in the same or different countries conduct projects leading toward the same or similar goals, interlaboratory test programs should be initiated between these peer laboratories. Examples of two types of such projects follow:

   a. SIP ambient air monitoring - An ambient air monitoring network is required by each of the 50 state air-pollution control agencies for criteria air pollutants* as part of State Implementation Plans (SIP). The EPA initiated an interlaboratory test program that includes government agencies (both state and local) that participate in SIP ambient air monitoring networks. An example of the type of EPA information obtained in the interlaboratory test programs is shown in Reference 2.

b. International ambient-air monitoring - Through the World Health Organization (WHO), international ambient air monitoring networks are being established.[3] In conjunction with this international monitoring network, interlaboratory test programs have been initiated. Interlaboratory test programs involving two or more countries are particularly important for air pollution measurement systems in which the data from one country will be used by another country in a decision-making process, such as establishing health effects criteria concentrations for specific air pollutants.

Factors affecting interlaboratory testing - Several factors may occur during the interlaboratory testing that may significantly affect the results. Some are administrative problems and some are technical problems related to the method.

How to obtain an unbiased treatment of reference samples is an important problem in an interlaboratory test program. Reference samples should be identified as routine samples with concentrations in the same range as routine samples. If the analyst recognizes samples as a reference sample from a referee laboratory, extra care will probably be taken in the analysis and the results may not be representative of routine analysis. However, unbiased treatment of reference samples by the analyst is an administrative problem the referee laboratory has little or no control over, except to point this problem out to the management of the participating laboratories.

---

* Sulfur dioxide, hydrocarbons, total oxidants, carbon monoxide, nitrogen dioxide and total suspended particulates.

Stability of the reference samples is an important technical consideration. The referee laboratory should determine the net effect of storage and transportation on the reference samples. A technique called "pigeon sampling" is sometimes used. A set of reference samples is prepared and a set of aliquots is analyzed. Another set of aliquots is shipped to a designated point and returned unopened to the referee laboratory. The aliquots are analyzed, and the results are compared with those obtained for aliquots before shipment.

The interpretation of written procedures for the interlaboratory testing is an important technical consideration. The interlaboratory testing procedures should be clear and complete:

1. So that all participating laboratories condition the reference samples when appropriate and control other parameters that may affect the final result.

2. So that all laboratory analyses are completed on a schedule that allows for the calculation of repeatability (different-day precision) and replicability (same-time precision), when these results are requested as part of the interlaboratory testing program.

3. So that all reported laboratory results are calculated correctly and in the proper metric units.

Collaborative testing - A special form of interlaboratory testing is collaborative testing. Laboratories participating in collaborative tests use the same measurement method procedure. Commonly, the coordinating laboratory answers questions from the participating laboratories, but step-by-step guidance is not provided by the coordinating laboratory at the time the method is used for the collaborative test. Collaborative testing is commonly used to measure the total performance of a measurement method, both sample collection and analysis, in terms of bias (accuracy) and precision (reproducibility, repeatability, and replicability).

Collaborative testing is an integral part of the EPA method standardization program for air pollution measurements of ambient air and source emissions. A listing of EPA collaborative study reports that are available is shown in Figure 1.4.15.1. Typically, an EPA collaborative study consists of the following steps:

METHODOLOGY STUDIES

A. AMBIENT AIR

Report Title	Reference Number
Collaborative Study of Reference Method for the Determination of Suspended Particulates in the Atmosphere (High-Volume Method).	PB 205-891, June 1971 (a)
Collaborative Study of Reference Method for Determination of Sulfur Dioxide in the Atmosphere (Pararosaniline Method).	PB 205-893, Sept. 1971 (a)
Collaborative Study of Reference Method for Determination of Sulfur Dioxide in the Atmosphere (Pararosaniline Method) (24-Hour Sampling).	EPA-650/4-74-027, 12/73 (b)
Collaborative Study of Reference Method for the Continuous Measurement of Carbon Monoxide in the Atmosphere (Non-Dispersive Infrared Spectrometry).	PB 211-265, 5/72 (a)
An Evaluation of the Arsenite Procedure for the Determination of Nitrogen Dioxide in Ambient Air.	EPA-650/4-74-048, 11/74 (b)
An Evaluation of the TGS-ANSA Procedure for the Determination of Nitrogen Dioxide in Ambient Air.	EPA-650/4-74-047, 11/74 (b)
Evaluation of Triethanolamine Procedure for Determination of Nitrogen Dioxide in Ambient Air.	EPA-650/4-74-031, 7/74 (b)
Collaborative Testing of Methods for Measurement of NO_2 in Ambient Air. Volume I - Report of Testing (Sodium Arsenite Procedure).	EPA-650/4-74-019a, 6/74 (b)
Collaborative Test of the TGS-ANSA Method for Measurement of Nitrogen Dioxide in Ambient Air.	EPA-650/4-74-046, 9/74 (b)
Survey of Users of the EPA - Reference Method for Measurement of Non-Methane Hydrocarbons in Ambient Air.	EPA-650/4-74-068, 12/74 (b)
Collaborative Test of the Continuous Colorimetric Method for Measurement of Nitrogen Dioxide in Ambient Air.	EPA-650/4-75-011, 2/75 (b)
Collaborative Test of the Chemiluminescent Method for Measurement of NO_2 in Ambient Air.	EPA-650/4-75-013, 2/75 (b)
Collaborative Study of Reference Method for Measurement of Photochemical Oxidants in the Atmosphere (Ozone-Ethylene Chemiluminescent Method).	EPA-650/4-75-016, 2/75 (b)
Evaluation of Effects of NO, CO_2, and Sampling Flow Rate on Arsenite Procedure for Measurement of NO_2 in Ambient Air.	EPA-650/4-75-019, 4/75 (b)
Evaluation of Gas Phase Titration Technique as Used for Calibration of Nitrogen Dioxide Chemiluminescence Analyzers.	EPA-650/4-75-021, 4/75 (b)
Evaluation of a Continuous Colorimetric Method for Measurement of Nitrogen Dioxide in Ambient Air.	EPA-650/4-74-022, 4/75 (b)

Figure 1.4.15.1 EPA method evaluations and collaborative studies on ambient-air and stationary source specific methods.

B. STATIONARY SOURCES

Report Title	Reference Number	
Collaborative Study of Method for the Determination of Nitrogen Oxide Emissions from Stationary Sources (Fossil Fuel-Fired Steam Generators).	EPA-650/4-74-025, 10/73	(b)
Collaborative Study of Method for the Determination of Sulfur Dioxide Emissions from Stationary Sources (Fossil Fuel-Fired Steam Generators).	EPA-650/4-74-024, 12/73	(b)
Collaborative Study of Method for the Determination of Particulate Matter Emissions from Stationary Sources (Fossil Fuel-Fired Steam Generators).	EPA-650/4-74-021, 6/74	(b)
Collaborative Study of Method for the Determination of Particulate Matter Emissions from Stationary Sources (Municipal Incinerators).	EPA-650/4-74-022, 7/74	(b)
Collaborative Study of Method for the Determination of Nitrogen Oxide Emissions from Stationary Sources (Nitric Acid Plants).	EPA-650/4-74-028, 5/74	(b)
The Collaborative Study of EPA Methods 5, 6 and 7 in Fossil Fuel-Fired Steam Generators - Final Report.	EPA-650/4-74-013, 5/74	(b)
Laboratory and Field Evaluations of EPA Methods 2, 6 and 7.	EPA-650/4-74-039,10/73	(b)
Collaborative Study of Method for the Determination of Particulate Matter Emissions from Stationary Sources (Portland Cement Plants).	EPA-650/4-74-029	(b)
Collaborative Study of Method for the Determination of Stack Gas Velocity and Volumetric Flow Rate in Conjunction with Method 5.	EPA-650/4-74-033, 9/74	(b)
Collaborative Study of Method for Stack Gas Analysis and Determination of Moisture Fraction with Use of Method 5.	EPA-650/4-74-026, 6/74	(b)
Collaborative Study of Method 104 - Reference Method for Determination of Beryllium Emissions from Stationary Sources.	EPA-650/4-74-023, 6/74	(b)
Collaborative Study of Method 10 - Reference Method for Determination of Carbon Monoxide Emissions from Stationary Sources - Report of Testing.	EPA-650/4-75-001, 1/75	(b)
Survey of Manual Methods of Measurements of Asbestos, Beryllium, Lead, Cadmium, Selenium, and Mercury in Stationary Source Emissions.	EPA-650/4-74-015, 9/73	(b)
Collaborative Study of Method for the Determination of Sulfuric Acid Mist and Sulfur Dioxide Emissions from Stationary Sources.	EPA-650/4-75-003, 11/74	(b)
Evaluation and Collaborative Study of Method for Visual Determination of Opacity of Emissions from Stationary Sources.	EPA-650/4-75-009, 1/75	(b)

(a) Available from the National Technical Information Service, Department of Commerce, Port Royal Road, Springfield Virginia 22151.

(b) Available from the Office of Forms and Publications, Environmental Protection Agency, Highway 70 Warehouse Facility Durham, North Carolina 27703.

Figure 1.4.15.1 (continued). EPA method evaluations and collaborative studies on ambient-air and source specific methods.

1. Ruggedness testing - The coordinating laboratory conducts a series of tests to determine the sensitivity (hopefully, to confirm the insensitivity) of the total measurement system, both sample collection and analysis, to variations of certain parameters suspected to affect the measurement system. The method write-up is reviewed and modified to improve clarity and completeness, particularly with regard to those measurement parameters determined from the ruggedness testing to be sensitive.

2. Collaborative testing - After ruggedness testing the coordinating laboratory arranges for a necessary number of participants to obtain meaningful statistics from the results. Normally 10 or more participants is desired. Each laboratory provides its own sample collection and analysis equipment. The test program is designed to determine measurement method bias (accuracy) and precision (reproducibility, repeatability, and replicability). Sample collection and analysis are performed on both reference materials and/or gases and on the actual environment for which the method is intended, namely, the ambient air for air methods and specific-source emissions for source methods.

Further details on both interlaboratory and collaborative testing may be found in Appendix K, - "Interlaboratory Tests."

Intralaboratory testing - The purpose of intralaboratory testing programs is to identify the sources of measurement method error and to estimate their bias (accuracy) and variability (repeatability and replicability). For manual measurement methods, bias and variability are determined separately for sample collection and analysis and are combined for determination of total method bias and variability. For continuous methods, total method bias and variability are determined directly. Some of the potential error sources are the operator or analyst, equipment, the calibration, and the operating conditions. The results may be analyzed by making comparisons against each other and/or against reference standards. Intralaboratory quality assurance procedures include techniques discussed in each of Sections 1.4.12, 1.4.13, 1.4.16, and 1.4.17. Operator or analyst proficiency is an additional consideration for intralaboratory testing.

Whereas many of the techniques employed in the conduct of an interlaboratory testing program are applicable in a modified form to intralaboratory testing, there are additional problems related to in-house proficiency testing of operators and analysts.

The major problems with designing a program to audit the analyst's proficiency are concerned with the following:

a. What kinds of samples to use.

b. How to prepare and introduce samples into the run without the analyst's knowledge.

c. How often to check the analyst's proficiency.

The problems and suggested solutions or criteria for decision are given in Figure 1.4.15.2.

1.4.15.3 REFERENCES

1. Pooler, F. The St. Louis Regional Air Pollution Study: A Coherent Effort Toward Improved Air Quality Simulation Models. Presented at the Summer Computer Simulation Conference, Houston, Texas. July 1974. Copies available from RAPS, EPA, Research Triangle Park, North Carolina 27711.

2. Bromberg, S.M., Akland, G.G., and Puzak, J.C. Survey of Laboratory Performance Analysis of Simulated Ambient SO_2 Bubbler Samples. Journal of the Air Pollution Control Association 24, 11. November 1974.

3. WHO International Air Pollution Monitoring Network--Data User's Guide. EP/72.6. June 1972. Available from Division of Environmental Health, WHO, 1211 Geneva 27, Switzerland.

BIBLIOGRAPHY

1. Youden, W. J., and Steiner, E.A. Statistical Manual of the Association of Official Analytical Chemists. (Statistical Techniques for Collaborative Tests - Youden; Planning and Analysis of Results of Collaborative Tests - Steiner), The Association of Official Analytical Chemists, P.O. Box 540, Benjamin Franklin Station, Washington, D.C. 20044. 1975.

Problem	Solutions and decision criteria
Kinds of samples	1. Use replicate samples of unknowns or reference standards. 2. Consider cost of samples. 3. Samples must be exposed by the analyst to same preparatory steps as are normal unknown samples.
Introducing the sample	1. Samples should have same labels and appearance as unknowns. 2. Because checking periods should not be obvious, supervisor and analyst should overlap the process of logging in samples. 3. Supervisor can place knowns or replicates into the system occasionally. 4. Save an aliquot from one day for analysis by another analyst. This technique can be used to detect bias.
Frequency of checking performance	1. Consider degree of automation. 2. Consider total method precision. 3. Consider analyst's training, attitude, and performance record.

Figure 1.4.15.2 Problems in assessing analyst proficiency.

1.4.16 AUDIT PROCEDURES

1.4.16.1 ABSTRACT

1. Performance audits are independent checks made by the supervisor or auditor to evaluate the quality of data produced by the total measurement system (sample collection, sample analysis and data processing). Performance audits are normally a quantitative appraisal of quality.

2. A system audit is an on-site inspection and review of the quality assurance system used for the total measurement system (sample collection, sample analysis, data processing, etc.). Quality assurance plans discussed in Section 1.4.23 should be used as the basis for conducting a system audit. System audits are normally a qualitative appraisal of quality.

1.4.16.2 DISCUSSION

Performance audit - Performance audits refer to independent checks made by the supervisor or auditor to evaluate the quality of data produced by the total sampling and analysis system. Performance audits generally are categorized as follows:

1. Sampling audits.
2. Analysis audits.
3. Data processing audits.

These audits are performed independent of and in addition to normal quality control checks by the operator/analyst. Independence can be achieved by having the audit made by a different operator/analyst from the one conducting the routine measurements or, in the case of sampling or analysis, by the introduction of audit control standards into the sampling or analysis system and the subsequent plotting of results on control charts by the supervisor. The use of audit control standards should be applied without the knowledge of the operator/analyst, if possible, to insure that recorded results reflect normal operating conditions.

Performance audits made by a different operator/analyst from the one conducting the routine measurement may be conducted in several ways. The following are examples of the most common type of audits.

1. Sampling audit - The auditor uses a separate set of calibrated flowmeters and reference standards to check the sample collection system:

   a. Flow rate devices.

   b. Instrument calibration.

   c. Instrument calibration gases, when applicable.

2. Analysis audits - The auditor is commonly provided a portion or aliquot of several routine samples for analysis.

3. Data processing audits - Data reporting commonly involves a spot-check on calculations and data validation may be checked by inserting in the data processing system a dummy set of raw data followed by review of these validated data.

A major problem in audit design is determining the auditing frequency and the lot size (number of samples required to estimate population concentration with a specified percentage confidence).

After the auditing schedule has been implemented, the audit results should be plotted on control charts to give the supervisor a visual picture of changes in the performance so that corrective actions may be taken when necessary. A discussion of recommended control chart types, instructions on construction, and criteria for interpretation are included in Appendix H, "Control Charts."

System audit - A system audit is an on-site inspection and review of the quality assurance system used for the total measurement system (sample collection, sample analysis, data processing, etc.) for each monitoring sensor. Whereas performance audits are a quantitative appraisal, system audits are normally a qualitative appraisal.

The quality assurance plans for projects and programs discussed in Section 1.4.23 should be used as the basis for conducting a system audit. For convenience, some of the items recommended for a QA plan are repeated here:

1. Organization and responsibility - Is the quality assurance organization operational?

2. Sample collection - Are written sample-collection procedures available and are these followed as written?

3. Sample analysis - Are written analysis procedures available and are these followed as written?

4. Data validation - Is a list of criteria for data validation available and is it used?

5. Calibration - Are written calibration procedures available and are these followed as written? In addition, a review should be made of procedures used to establish traceability of calibration standards to standards of higher accuracy and of the calibration schedule and data by measurement sensor.

6. Audits - Are control charts for performance audits reviewed?

7. Interlaboratory tests - Are results from interlaboratory testing reviewed?

8. Preventive maintenance - Is the preventive maintenance schedule being followed as recommended in the QA plan?

Other items for consideration in a system audit are shown in Section 2.0 of Volume II, Ambient-Air Specific Methods and Section 3.0 of Volume III, Stationary-Source Specific Methods.

A system audit may be made at any time during the life of a project but is normally conducted before or just after monitoring has been initiated.

1.4.17 DATA VALIDATION

1.4.17.1 ABSTRACT

Data validation can be accomplished by several methods. Validation can be manual or computerized.

1. Data validation is the process whereby data are filtered and accepted or rejected based on a set of criteria. Validation is performed to isolate spurious values since values are not automatically rejected. Records of invalid data should be maintained.

2. Validation methods can include review by supervisory personnel as well as application of validation criteria by computer. Criteria depend on the types of data and on the purpose of the measurement.

3. A number of statistical techniques are useful. Periodic checking of manually reduced data values is important. Important statistical techniques are:

   a. Acceptance sampling.

   b. Sequential analysis.

4. Data validation procedures and specific criteria for the EPA National Aerometric Data Bank (NADB) are given in order to illustrate important areas of concern which should be considered. These areas include:

   a. Screening data for representativeness; instrument averaging time; sampling program duration; and comparability with other reported data.

b. Providing criteria for: completeness of data; use of accuracy and precision data; handling of data reported as below the minimum detectable limits; and handling of data reported with negative values.

1.4.17.2 DISCUSSION

Data validation is the process whereby data are filtered and accepted or rejected based on a set of criteria. This involves a critical review of a body of data in order to isolate and locate spurious values. It may involve only a cursory scan to detect extreme values or a detailed evaluation requiring the use of a computer. In either situation, when a spurious value is located, it is not immediately rejected. Each questionable value must be checked for validity. Records of values that are either judged invalid or are otherwise suspicious should be maintained. These records are, among other things, a useful source of information for judging data quality.

Methods of data validation

1. Manual - Both the analyst and the laboratory supervisor should inspect integrated air quality monitoring data and emission source testing data. At regular intervals, daily or weekly, results should be scanned for questionable values. This type of validation is most sensitive to extreme values, i.e., either unusually high or low concentrations. These are sometimes called outliers.

The criteria for determining an extreme value are derived from prior data obtained at the particular sampling site (or a similar site if no previous data are available for a site). The data used to determine extremes may be the minimum and maximum concentrations for all prior data from a site. The decision criteria might also be based on minimum and maximum for each season, each month, or each day. See Appendix F for a further discussion on outliers.

The time spent checking data that have been manually reduced by technicians depends on the time available and on the demonstrated abilities of the technicians to follow instructions. The method instructions in Volume II of this Handbook suggest an audit level of 7 percent, i.e., checking 7 out of every 100 values. This audit level is somewhat arbitrary, and it would be expected to be changed as more experience with the method is gained. Criteria for determining acceptability of data based on the difference between audit values and original data values are also suggested in the audit sections of Volume II.

2. Computerized Techniques - A computer can be used not only to store and retrieve data but also to validate data. The system for checking extreme values in manual techniques also applies here. The criteria for extreme values can be refined to be specific for individual hours during the day. With this procedure, an hourly average

concentration for carbon monoxide of 15 ppm may not be considered an extreme value for 8:00 a.m. but could be tagged as questionable if it appeared at 2:00 a.m.

Another indication of possible spurious data is a large difference in concentrations reported for two successive time intervals. The difference in concentrations, which might be considered excessive, may vary from one pollutant to another and quite possibly may vary from one sampling location to another for the same pollutant. Ideally this difference in concentration is determined through a statistical analysis of historical data. For example, it may be determined that a difference of 0.05 ppm in the SO_2 concentration for successive hourly averages occurs rarely (less than 5 percent of the time). But at the same station the hourly average CO concentration may change by as much as 10 ppm. The criteria for what constitutes an excessive change may also be linked to time of day and pollutant relationships, e.g., high concentrations of SO_2 and 0_3 can not co-exist and these data should be considered suspect.

The validation criteria for any data set should ultimately be determined by the objectives for collecting the data. The extent of the decision elements to be used in data validation can not be defined for the general case. Rather, the validation criteria should be tailored along the lines suggested above for varying types of air monitoring networks.

Tools used in data validation - There are a number of statistical tools that can be used in data validation.

1. General - Often the output from a continuous air monitoring device is an analog trace on a strip chart. Usually the strip charts are cut at weekly intervals and are sent to the data-handling staff for interpretation. A technician may estimate by inspection the hourly average pollutant concentrations and convert the analog percent of scale to other units, e.g., ppm.

Reading strip charts is a tedious job subject to varying degrees of error. A procedure for maintaining a desirable quality for data manually reduced from strip charts is important. One procedure for checking the validity of the data reduced by a technician is to have another technician or the supervisor check the data. Because the values have been taken from the strip chart by visual inspection, some difference in the values derived by two individuals can be expected. When the difference exceeds a specified amount and the initial reading has been determined to be incorrect, an error should be noted. If the number of errors exceeds a predetermined number, all data for the strip chart are rejected and the chart is read again by a technician other than the one who initially read the chart. The question of how many values to check can be answered by applying acceptance sampling techniques.

2. Application of Acceptance Sampling Techniques - Acceptance sampling can be applied to data validation to determine the number of data items (individual values on a strip chart) that need to be checked to determine with a given probability that all the data items are acceptable. The supervisor wants to know, without checking every data value, if a defined error level has been exceeded. From each strip chart with N data values, the supervisor can randomly inspect n data values. If the number of erroneous values is less than or equal to c, the rejection criteria, the values for the strip chart are accepted. If the number of errors is greater than c, the values for the strip chart are rejected, and another technician is asked to read the chart.

3. Sequential Analysis Test Procedure - The typical approach used in performing a statistical test of hypothesis requires the collection of a sample of a fixed size. A statistic is then computed from the sample data and compared with some critical values for that statistic. A decision is then made to accept the hypothesis (H_o) or to accept some alternative hypothesis (H_1). With such a procedure it is necessary to collect the specified sample of observations regardless of the results that may be obtained from the first few observations.

A procedure called sequential analysis requires that a decision be made after each observation or group of observations. This procedure has the advantage that, on the average, a decision can be reached with fewer observations than a fixed sample size requires.

Data validation procedures and criteria for EPA's National Aerometric Data Bank - Data validation procedures and specific criteria for the EPA National Aerometric Data Bank (NADB) are given as an example to illustrate important characteristics of data validation that should be considered. Use of procedures similar to those developed for NADB will facilitate data interpretations and comparisons among investigators. The example shown here is for ambient air monitoring data. A more detailed discussion of data handling and validation procedures for NADB is described in the literature.[1]

1. Scope of NADB - The NADB is a computer storage and retrieval system for aerometric data collected by Federal, State, and local agencies.

2. Screening Criteria - In order to draw correct conclusions from the data, validity checks are built into the data handling system. The data must meet predetermined standards with respect to representativeness, instrument averaging time, duration of sampling, and comparability before they are incorporated into NADB. A discussion of each criterion follows:

a. Representativeness. Data from each monitoring site should characterize ambient levels in an area or neighborhood. For example, a daily average of carbon monoxide calculated from values collected only during the morning rush hour would hardly reflect the true daily averages. The data must be relatively complete over the time interval of interest (for example, day, season, or year) so that such biases can be avoided.

b. Instrument Averaging Time. The data must represent a sample interval of 1 hour or more. Thus, no more than 24 values per day per pollutant are stored. Data for intervals of less than 1 hour are converted to hourly averages before storage.

c. Duration. The data must be collected over a time period of no less than 3 consecutive months so that at least quarterly summary statistics can be calculated.

d. Comparability. Aerometric data must be maintained in consistent units to permit data submitted by various agencies to be combined into nationwide summaries and evaluation reports. The data must have been acquired by application of standard methodologies.

These four criteria are pertinent to developing meaningful information from the data collected from any monitoring network.

3. Criteria for Completeness. The raw data entering the NADB are checked for completeness (representativeness).

With continuous measurement, the criterion for completeness is that at least 75 percent of the total possible number of observations be present. Figure 1.4.17.1 presents the number of observations required by the NADB before summary results are calculated.

The data within the NADB resulting from intermittent sampling are summarized only if there are at least five samples per quarter. An additional stipulation is that if a month contains no samples each of the other 2 months in the quarter must contain at least 2 samples. Any other distribution of samples over the quarter is acceptable. This is a minimal criterion based on a random biweekly sampling schedule. A more stringent criterion should be applied when the sampling schedule is every third or sixth day.

4. Criteria for Accuracy and Precision - Accuracy and precision data reported with aerometric data are not used to validate data before entry into NADB but are used to interpret the data.

5. Criteria for Handling Data Values below Minimum Detectable Limits - Concentrations below the limit of detection of the instruments employed result in the problem of determining how to report such values so that summary statistics can be calculated. The choice of data values is complicated by the fact that zero, the most likely value to be supplied, cannot be used, especially if geometric parameters are to be calculated.

Continuous measurement criteria for completeness

Time interval	Minimum number of observations
3-hour running average	3 consecutive hourly observations
8-hour running average	6 hourly observations
24-hour	18 hourly observations
Monthly	21 daily averages
Quarterly	3 consecutive monthly averages
Yearly	9 monthly averages with at least two monthly averages per quarter

Figure 1.4.17.1 Criteria for completeness for continuous ambient air monitors for NADB

This problem is handled by inserting a constant approximately equal to one-half the minimum detectable limit for each method and analysis technique. This value was chosen after examining the lower end of the cumulative distributions for the various pollutants. Seldom did the log-normal distribution (the distribution most often applied to air pollution data) accurately describe this portion of the data. This may be due in part to the existence of a background level for each pollutant. Use of the midpoint between zero and the detectable limit as the substitute value for concentration levels below the detection threshold therefore seems reasonable. In order to permit consistency from year to year, one minimum detectable value is used for each pollutant even if the minimum detectable limit is changed, unless there is a change by an order of magnitude. Figure 1.4.17.2 provides an example of the current minimum detectable limits as used by the NADB for each pollutant and the value to be inserted for each value below the minimum detectable limit. These minimum detectable limits are reviewed periodically and changed as required. Each laboratory should determine its own set of minimum detectable limits based on its own analytical techniques and instruments used to generate pollutant information.

One additional point should be mentioned concerning the use of substituted values for values below the threshold of

Pollutant	Method of Analysis	Minimum detectable limit, $\mu g/m^3$	Substitution factor, $\mu g/m^3$
Suspended particulate	Gravimetric	1.0	0.5
Fluoride	Specific ion electrode	0.05	0.025
Nitrate	Reduction-diazo	0.05	0.025
Sulfate	Colorimetric	0.5	0.025
Carbon monoxide	Nondispersive infrared	573	286
Sulfur dioxide	Pararosaniline, bubbler	5	2.5
Sulfur dioxide	Pararosaniline, continuous	26	13
Ozone	Chemiluminescence	20	10
Total oxidant	Colorimetric	20	10
Nitrogen dioxide	Arsenite, bubbler	5	2.5
Nitrogen dioxide	TGS-ANSA, bubbler	5	2.5
Nitrogen dioxide	Chemiluminescence, continuous	9	4.5
Nitrogen dioxide	Colorimetric, continuous	9	4.5

Figure 1.4.17.2 Minimum detectable limits, by pollutant and method of analysis, used in NADB in 1975

the method. When more than 25 percent of the measured levels are below the minimum detectable quantity, no statistics are computed from the data. This constraint guards against the possibility of biasing the computed statistics. Furthermore, at least 50 percent of the measurements in a set of data must be above the minimum detectable concentration before a frequency distribution of the values is prepared.

6. Criteria for Handling Data with Negative Values - For the purpose of generating true pollutant values, negative pollutant concentrations imply that there is not enough of the pollutant present in order for the instrument to detect above the noise limit specified for the instrument. When negative values occur, they should be regarded as being below the detection limit for the method and treated in the same manner, i.e., assigned a value one-half the minimum detectable limit.

1.4.17.3 REFERENCES

1. Nehls, G.J., and Akland, G.G. Procedures for Handling Aerometric Data. Journal of the Air Pollution Control Association. 23(3):180-184, March 1973.

1.4.18 STATISTICAL ANALYSIS OF DATA

1.4.18.1 ABSTRACT

A number of statistical tools and techniques are described in the appendices. The appendices are organized in part by functional or application area rather than by statistical nomenclature. For example, Appendix J concerns the subject of calibration; however, least squares or regression analysis, a useful tool for determining calibration curves, can also be used for estimating the resulting precision of the reported pollutant concentration from a specific analyzer reading. The statistical tools should be used with discretion.

A glossary of major statistical terms is included as Appendix A. Appendix B includes symbol definitions used throughout the remaining appendices.

1.4.18.2 DISCUSSION

Summary statistics - Summary statistics such as the mean and the standard deviation are used to simplify the presentation of data and at the same time to summarize essential characteristics. Appendix C includes a discussion of summary statistics.

Frequency distributions - Frequency distributions such as normal and log-normal distributions are used to present relatively large data sets, such as the daily concentrations of suspended particulates in ambient air. Appendix D discusses frequency distributions.

Estimation and testing procedures - Statistical estimation and testing procedures are used to make inferences concerning the conceptual population of measurements made under the same conditions based on a small sample of data. An example would be the estimation of the average pH of a large number (population) of filters based on a sample (lot) of pH readings for seven filters. Appendix E discusses estimation and testing procedures.

Outliers - Outliers, i.e., unusually large or small values, are identified by appropriate statistical tests for outliers. These statistical tests are useful, for example, in identifying gross errors in data handling procedures. Appendix F is a treatment of outliers.

Audit data - Methods for treating performance audit data and for presenting the results in terms of bias and precision are included in Appendix G.

Control charts - Techniques for selecting the type of control chart, for determining the limits, and for interpreting plotted results are presented in Appendix H.

Sampling - Sampling techniques apply to many phases of a quality assurance program. Methods for selecting a random sample, as well as procedures for acceptance sampling, are discussed in Appendix I.

Calibration - Calibration procedures represent one of the critical sources of measurement error. Appendix J is a discussion of calibration procedures. Techniques to indicate when a new multipoint calibration is to be conducted should be employed.

Replication, repeatability, and reproducibility tests - The identification of sources of measurement error within and among laboratories is one of the important functions of the Quality Assurance Coordinator. Programs for doing this are discussed in Appendix K.

Reliability and maintainability - As measurement systems become more complex, system reliability becomes an increasingly important parameter in determining the completeness and accuracy of the results. Reliability is discussed in Appendix L.

Tables - Statistical tables are given in Appendix M.

1.4.18.3 BIBLIOGRAPHY

1. Burr, I.W. Engineering Statistics and Quality Control. McGraw-Hill, New York. 1953.

2. Duncan, A.J. Quality Control and Industrial Statistics. 3rd Ed. Richard D. Irwin, Inc., Homewood, Illinois. 1965.

3. Juran, J.M., (ed.) Quality Control Handbook. 2nd Ed. McGraw-Hill, New York. 1962. Section 13. pp. 1-127.

4. Snedecor, G.W., and Cochran, W.G. Statistical Methods. 6th Ed. Iowa State College Press, Ames, Iowa. 1967.

1.4.19 CONFIGURATION CONTROL

1.4.19.1 ABSTRACT

1. Configuration control is used to record changes in air pollution measurement method equipment and the physical arrangement of this equipment in the monitoring system.

2. Configuration control may be grouped into two types depending on the purpose:

   a. Provides history (record) of changes during the life of the monitoring project.

   b. Provides design and operation data on the first monitoring instrument or system when multiples are planned. This information is commonly obtained by a First Article Configuration Inspection (FACI). An example of a FACI is shown for a major EPA monitoring network in the discussion portion.

3. Configuration control record procedures are the same as those used for document control (Section 1.4.1).

1.4.19.2 DISCUSSION

Difference between Configuration and Document Control - Document control, described in Section 1.4.1, is used to make sure all personnel on a monitoring project are using the same and most current written procedures for sampling, analysis, calibration, data collection and reporting, auditing, etc. When revisions are made in these procedures, they should be documented as described in Section 1.4.1.

Similarly, a system is needed to record changes made in the equipment and/or physical arrangement of this equipment in the monitoring system that are not included as part of document control. This system is called configuration control.

Types of Configuration Control - Configuration control may be grouped into two types, depending on the intended purpose of the information.

In the first type, a history of changes is maintained throughout the life of the monitoring project. This history is valuable during problem-solving investigations that may occur either during the project life or long after the project has been completed. Subtle changes in the equipment used in the monitoring system may have significant effects on the measured pollutant concentrations. Such equipment changes would normally not appear under document control on the procedure used for sampling, analysis, etc. By way of example, some of these changes might include:

1. Replacement of monitoring instrument or component part with a different model type (equipment change).

2. Replacement of filter used to remove particulates prior to instrumental gaseous-pollutant analysis with a different filter type (equipment change).

3. Relocating an air pollution sampler to a different spot at the sampling site (rearrangement of same equipment).

Each project officer must decide the scope of configuration control that should be applied to his project.

The second type of configuration control is used to provide information on engineering design and operation on the first monitoring instrument or station when multiples are planned. This information is commonly obtained and documented by a First Article Configuration Inspection (FACI). The FACI is most important for large complex monitoring projects, particularly when pollutant sensor outputs are stored on-site or transmitted to a central facility for computer storage. Purchase contracts that involve multiple instrument systems of identical design and/or monitoring stations of identical design should require a FACI as part of the contract.

By way of example, the FACI required as part of the contract for the EPA Regional Air Monitoring System (RAMS) will be briefly described. The RAMS is a network of 25 monitoring sites in and around the St. Louis area, designed to collect ambient air and meteorological measurements for diffusion modeling and other purposes. When the first monitoring station was installed, a FACI was completed as required by the contract. The FACI covered the following:

1. Shelter system.

2. Gas analyzing system (sensor for ozone, nitrogen oxides, total hydrocarbons, carbon monoxide, and total sulfur.

3. Particulate sampling system (including sensor for light scattering).

4. Meteorological system (sensor for wind speed, wind direction, temperature and dew point).

5. Data acquisition system.

For each system, the FACI consisted of a physical inspection, a functional demonstration, and an operational test consistent with requirements in the contract. To facilitate and semi-formalize the exchange of information between EPA and the contractor during the FACI, "squawk sheets" were used. These sheets allowed discrepancies to be noted by EPA and were responded to by the contractor. An example of the RAMS squawk sheet is shown in Figure 1.4.19.1. The contractor prepared a formal response to all squawk sheets.

The procedures described in Section 1.4.1 for document control are also applicable for configuration control of hardware over the project life.

1.4.19.3 REFERENCES

1. Lowers, H.R. Quality Assurance in Configuration Management. Quality Progress. V(6):17-19, June 1972.

2. Covino, C.P., and Meghri, A.W. Quality Assurance Manual. Industrial Press, Inc., New York. 1967. pp. 31a, 31b, and 32a.

BIBLIOGRAPHY

1. Configuration Management. Air Force Systems Command Manual 375-1. 1960.

Squawk Title ______________________________

No. ______________
Date _____________
Author ____________

Squawk Description:

EPA
Coordinator _______

[EPA Comment]

Contractor Action/Response

Contractor Program
Engineer __________

[Contractor response to EPA comment]

Final Disposition

Contractor Program Engineer _________________________

EPA Project Officer _________________________________

Figure 1.4.19.1 RAMS - FACI Squawk Sheet

1.4.20 RELIABILITY

1.4.20.1 ABSTRACT

Reliability of an air pollution measurement system is defined as the probability that the system will perform its intended function for a prescribed period of time under the operating conditions specified, or, conversely, unreliability is the probability that a device will fail to perform as specified. Reliability is becoming increasingly important in air pollution measurement because of the increase in complexity and sophistication of sampling, analysis, automatic recording, and telemetering systems. Furthermore, data interpretation depends on data completeness (no breaks in the data set) for trend analysis. Generally, as equipment becomes more complicated, its probability of failure increases. In order to assure high reliability (i.e., completeness of data) the following should be considered:

1. Specify equipment reliability in contracts--select high reliability components.

2. Inspect and test incoming equipment for adherence to contract specifications--e.g., burn-in and/or demonstrated performance acceptance tests.

3. Control the operating environment that influences the reliability of the measurements.

4. Provide for adeqate training of personnel.

5. Provide preventive maintenance to reduce or minimize wear out failures.

6. Provide records of failures, analyze and use these data to initiate corrective actions, and predict reliability rates.

1.4.20.2 DISCUSSION

Reliability of an air pollution measurement system is defined as the probability that the system will perform its intended function for a prescribed period of time under the operating conditions specified, or, conversely, unreliability is the probability that a device will fail to perform as specified. Reliability is becoming increasingly important in air pollution measurement because of the increase in complexity and sophistication of sampling, analysis, automatic recording, and telemetering systems. Furthermore, data interpretation depends on data completeness (no breaks in the data set) for trend analysis. Generally, as equipment becomes more complicated, its probability of failure increases.

In order to assure high reliability (i.e., completeness of data), the following should be considered:

Specify equipment reliability requirements in contracts[1] - These requirements constitute a specification to be met by the manufactured product. This specification should consist of:

1. The product reliability definition, which includes:

    a. All functional requirements of the equipment.

b. Safety requirements.

c. Environmental conditions for the reliability demonstration tests.

2. Where applicable, give required reliability expressed as a minimum mean time between failures (MTBF).[2] The MTBF is the average time that the system performs its required function without failure. This may be expressed as hours, days, or number of monitoring periods. It is estimated by averaging the recorded times of successful system performance.

3. Required reliability demonstration tests.

a. Burn-in test for a specified time to debug the early failures resulting from poor manufacturing and assembly techniques.

b. Reliability demonstration tests such as given in the military specifications, MIL-STD-781B.[3]

Inspect and test incoming equipment for adherence to contract specifications

1. Quality control tests should be conducted to determine whether the product in question meets performance and design specifications at the time of testing.

2. Burn-in tests should be conducted for specified times where there is an indication of early failures.

3. Reliability demonstration tests should be conducted on a sample of equipments, testing until failure or for a specified time, to:

a. Verify adherence to specified reliability standards.

b. Generate data for product improvement.

c. Provide an estimate of product service life and reliability.

Control the operating conditions[4] - Environmental factors affecting performance or reliability may be natural, induced, or a combination of both.

1. Natural environmental factors are:
    a. Barometric pressure changes.
    b. Temperature.
    c. Particulate matter, such as sand, dust, insects, and fungus.
    d. Moisture, such as icing and salt spray.
2. Induced factors are:
    a. Temperature, self-generated or generated by adjacent or ancillary equipment.
    b. Dynamic stresses, such as shock vibration.
    c. Gaseous and particulate contamination, such as exhaust or combustion emissions.
3. Combined natural and induced conditions. Frequently, the stresses affecting an item result from a combination of one or more factors from both classes. Such combinations may intensify the stress, or the combined factors may tend to cancel out each other.

Provide for adequate training of personnel[5,6] - The implementation of a reliability assurance program requires a training program at both the operational and supervisory levels. At the operator level, instruction should be given in the collection of failure and maintenance data, in the maintenance function (both preventive and unscheduled maintenance or repair of the equipment), and in the control of operating conditions. This training can be accomplished by use of lectures, films, posters, and reliability information bulletins.

At the supervisory level, in addition to the above, training should be given in the analysis of reported data, program planning, testing, and demonstration procedures.

The reliability of the measurement system depends to a large extent on the training of the operator. The completeness of the data, as measured by the proportion of valid data reported, is a function of both the reliability and maintainability.

Consider maintainability at time of purchase - Maintainability is the probability that the system will be returned to its operational state within a specified time after failure. For continuous air pollution monitoring instruments, maintainability is an important consideration during procurement, and in some cases should be included in the purchase contract. Maintainability items to consider at the time of procurement include:

1. Design factors.
    a. Number of moving parts.
    b. Number of highly stressed parts.
    c. Number of heat producing parts.
2. Ease of repair after failure has occurred.
3. Maintainability cost.
    a. Inventory of spare parts required.
    b. Amount of technician training required for repair.
    c. Factory service required.
    d. Service repair contract required.
    e. Estimated preventive maintenance required.

Provide preventive maintenance - In order to prevent or minimize the occurrence of wear out failure, the components of the system subject to wear out must be identified and a preventive maintenance schedule implemented. This aids in improving the completeness of the data. Maintenance can be performed during nonoperational times for noncontinuous monitoring equipments, resulting in no downtime. Replacement units must be employed in continuous monitoring systems in order to perform the maintenance while the system is performing its function. Downtime may also be scheduled.

Provide records of failure and maintenance; analyze and use to initiate corrective actions - Field reliability data should be collected in order to:

1. Provide information upon which to base reliability rate predictions.

2. Provide specific failure data for equipment improvement efforts.

3. Provide part of the information needed for corrective action recommendations.

A more complete discussion of reliability and maintainability is contained in Appendix L.

1.4.20.3 REFERENCES

1. Muench, J.O. A Complete Reliability Program. Proceedings Annual Reliability and Maintainability Symposium. Institute of Electrical and Electronic Engineers, Inc. 1972. pp. 20-23.

2. Enrick, N.L. Quality Control and Reliability. 6th ed. Industrial Press, Inc., New York, New York. 1972. Chapter 16, pp. 219-238.

3. MIL-STD-718B, Reliability Tests, Exponential Distribution. 1967. Department of Defense, Washington, D.C.

4. Haviland, R.P. Engineering Reliability and Long Life Design. D. Van Nostrand Co., Inc., Princeton, New Jersey. 1964. Chapter 10, pp. 149-168.

5. Bazovsky, I. Reliability Theory and Practice. Prentice-Hall, Englewood Cliffs, New Jersey. 1961. Chapter 9, pp. 76-84.

6. Juran, J.M. Quality Control Handbook. 2nd edition. McGraw-Hill, New York. 1962. Sec. 20, pp. 20-2 - 20-38 and Sec. 13, pp. 13-17 - 13-19.

1.4.21 QUALITY REPORTS TO MANAGEMENT

1.4.21.1 ABSTRACT

Quality facts to be reported should be defined and methods for summarizing and reporting data should be determined.

Quality facts usually reported are:

1. Percentage duplication or replication of determinations.
2. Instrument or equipment downtime.
3. Percentage voided samples versus total samples.
4. Quality cost in terms of prevention, appraisal, and correction costs.
5. System audit (on-site inspection) results.
6. Performance audit results.
7. Interlaboratory test results and, where appropriate, intralaboratory test results (precision and accuracy).
8. Status of solutions to major quality assurance problems.

The Quality Assurance Coordinator should consult with line staff to determine the necessary reporting requirements. Reports should be obtained from source documents. Reports should have a baseline for comparison and should be easy to interpret. Similar reports should be combined if possible.

1.4.21.2 DISCUSSION[1]

Quality assurance facts to be reported may vary widely from one air pollution control agency to another, depending on size and organization of the agencies. The quality reports listed in the abstract represent the kinds of information commonly needed by management.

The details or sources of such facts are obtained from other sections of the Handbook as follows:

	Facts	Source
1.	Percentage of duplication or replication of determinations.	Methods (Vols. 2 and 3)
2.	Instrument or equipment downtime.	1.4.20 (Vol. 1)
3.	Percentage of voided samples versus total samples.	Methods (Vols. 2 and 3) 1.4.17 (Vol. 1)
4.	Quality costs (prevention, appraisal, and correction costs).	1.4.14 (Vol. 1)
5.	System audit (on-site inspection) results.	1.4.16 (Vol. 1)
6.	Performance audit results.	1.4.16 (Vol. 1)
7.	Interlaboratory test results and, where appropriate, intralaboratory tests results (precision and accuracy).	1.4.15 (Vol. 1)
8.	Status of solutions to major quality problems.	1.4.13 (Vol. 1)

Principles of reporting to management - The following are general principles that should be considered:

1. In order to determine what periodic reports are needed or desired by management, the Quality Assurance

Coordinator should discuss with superiors and staff members what reports need to be published on a regular basis to provide information required for rational decision-making.

2. In order to minimize errors in transmission, translation, and interpretation, facts used in reports to management should be obtained from source documents when possible.

3. Facts should be presented in summary form. When not presented graphically, the report should provide abstracts of essential points.

4. Reports should compare current data with some base. The basic standard of comparison can be:

   a. Based on previous performance (historic).

   b. Based on engineering judgment.

5. The report should be understood at a glance. If the report is not presented graphically, essential points should be reduced to a single page, preferably double-spaced. An example of a graphic report is shown in Figure 1.4.21.1. Figure 1.4.14.1 for quality costs (in terms of prevention, appraisal, and corrective costs) is another example of graphic presentation.

Where applicable, quality control charts should be used to report results from performance audits, intralaboratory tests and, when performed on a recurring basis, interlaboratory test results.

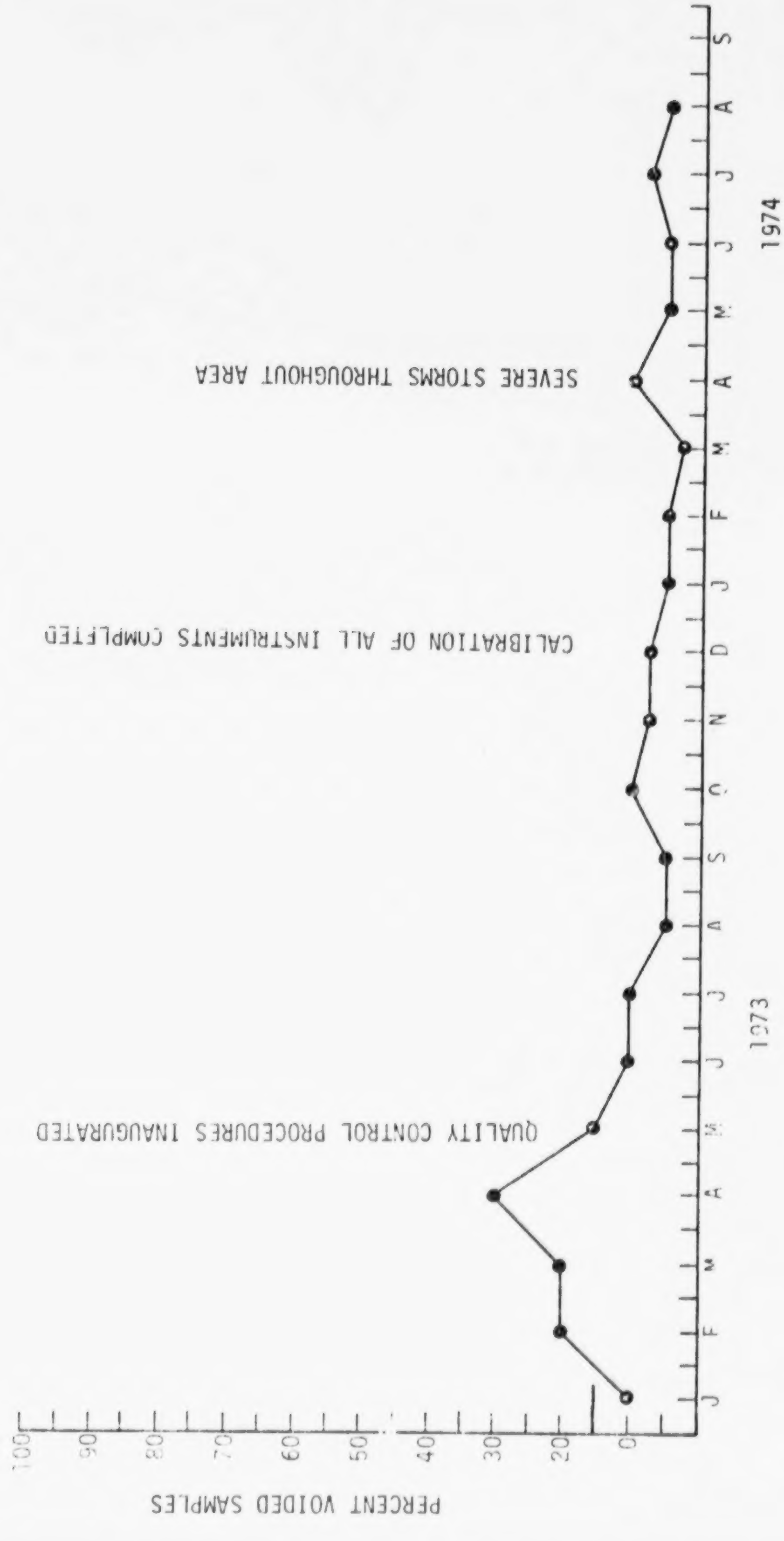


Figure 1.4.21.1 Graphic report to management.

1.4.21.3 REFERENCES

1. Juran, J.M., (ed.). Quality Control Handbook. 2nd ed. McGraw-Hill, New York. 1962. Sec. 12, pp. 26-32.

BIBLIOGRAPHY

1. Juran, J.M., and Gryna, F.M. Quality Planning and Analysis. McGraw-Hill, New York. 1970.

2. Spear, M.E. Charting Statistics. McGraw-Hill, New York. 1952.

1.4.22 QUALITY ASSURANCE MANUAL

1.4.22.1 ABSTRACT

1. A quality assurance manual is a manual of general requirements needed to assure quality. Specific requirements for each project are contained in a quality assurance plan.

2. Each state and local air pollution control agency and others involved in air program research and monitoring activities need a quality assurance manual. These organizations should use the "Elements of Quality Assurance" described in Section 1.4 of this Handbook as a guide in developing their own manuals.

1.4.22.2 DISCUSSION

Quality assurance manual content - A common practice for quality assurance in industry is for each company concerned with the design, development, and manufacturing of a product to prepare a quality assurance manual. This manual states the company's general philosophy and requirements with respect to quality assurance by functional activities that are associated with the quality of the product. The specific requirements for each product (or group of products) are contained in a quality assurance plan.

State and local air pollution control agencies and others need a quality assurance manual - Each state and local air pollution control agency and others involved in air program research and monitoring activities need a quality assurance manual. The manual should define management's position on quality assurance and provide guidelines that should be followed by subordinates in preparing quality assurance plans. These organizations should use the "Elements of Quality Assurance" described in Section 1.4 of this Handbook as a guide in developing their own manual.

1.4.23 QUALITY ASSURANCE PLANS FOR PROJECTS AND PROGRAMS

1.4.23.1 ABSTRACT

1. A quality assurance plan (QA plan) is specific quality assurance requirements prepared for a project or program for the elements discussed in Section 1.4 of this Handbook.

2. QA plans should be prepared for both inhouse and contract projects.

3. The depth of a QA plan is influenced primarily by:

   a. Intended use of the data.

   b. Whether the project is new or on-going.

4. An example format for a QA plan is given.

1.4.23.2 DISCUSSION

Definition of quality assurance plan (QA plan) - The specific requirements prepared by or for a project officer (single project) or a program manager* (group of similar projects) to ensure quality data from each air measurement system are called a QA plan.

Quality assurance plans for inhouse and contract projects - A QA plan should be prepared and implemented to include each air pollution measurement system in an inhouse or contract project. The plan should be prepared by the project officer for inhouse projects and for the project officer's review and approval for contracts. The Quality

---

* For simplicity the terms project officer and project will be used throughout the discussion.

Assurance Coordinator should be involved in the review of the QA plan prior to implementation. The plan should be based on project objectives and should be consistent with management's general requirements on quality assurance that are specified in the quality assurance manual (see Section 1.4.22).

Quality assurance elements in a QA plan - All quality assurance elements discussed in Section 1.4 should be considered in the preparation of the QA plan. However, different projects will require emphasis on different elements. Factors that will have the most influence on the QA plan design are:

1. Intended use of the data - Monitoring for enforcement or the establishment of air pollution criteria levels (normally for health effects) will result in closer scrutiny of the project from a quality assurance standpoint than projects in which legal actions may not be involved; e.g., research on the development of a new monitoring instrument.

2. New project versus improvement in an on-going project - Improving quality assurance on a current project normally means that the operators are on the job and the air measurement equipment is in operation. For on-going projects, the quality assurance elements that are particularly beneficial to consider in the QA plan are:

    a. Organization (Section 1.4.3); namely, the project quality organization and responsibility.

b. Sample collection (Section 1.4.8), includes sampling handling and storage requirements. Include a copy of all procedures, or if standard procedures are used, a reference to these will suffice.

c. Sample analysis (Section 1.4.9). Include a copy of all procedures, or if standard procedures are used, a reference to these will suffice.

d. Data reporting (Section 1.4.10).

e. Data validation (Section 1.4.17), including specific criteria that are to be applied for validation of data.

f. Audit procedures (Section 1.4.16) including specific performance audits that are to be performed during routine project operations.

g. Calibration (Section 1.4.12), including description of procedure for establishing traceability, description of calibration standards, and a schedule for calibration. Include a copy of all calibration procedures, or if standard procedures are used, a reference to these will suffice.

h. Preventive maintenance (Section 1.4.7), including a schedule for maintenance.

i. Interlaboratory tests (Section 1.4.15), including a list of any planned or anticipated participation in interlaboratory tests.

j. Quality reports to management (Section 1.4.21), including the information content of the reports and the frequency of reporting.

Project officers or managers for new projects will need to consider all elements mentioned previously for on-going projects and, in addition, should probably pay special attention also to the following areas:

a. Pretest preparation (Section 1.4.6). For new monitoring projects, a preliminary visit to select sampling sites for ambient air monitoring and source emission monitoring will normally be required. Criteria or factors used to select the sample collection sites should be listed.

b. Training (Section 1.4.5). Commonly, new projects may result in the hiring of new personnel or the assignment of current personnel to different work tasks for which training may be required. Anticipated training should be listed.

c. Procurement quality control (Section 1.4.11). A new project may mean procurement of new equipment and reagents. Equipment and reagents that should be placed under procurement quality control should be listed and, if available, acceptance limits should also be included.

d. Reliability (Section 1.4.20). When equipment reliability is important, the equipment should be listed and reliability limits provided.

e. Configuration control (Section 1.4.19). Any plans for configuration control on large projects should be summarized.

Quality assurance plan document control and distribution - The preparation of a QA plan requires that the project officer consider all quality assurance elements listed previously and plan for those elements that will have the most quality impact in the project. The result of this planning is the QA plan *per se*. The QA plan represents the project officer's best judgment on the quality assurance level required for the project. Therefore, the QA plan should be readily available to all project personnel with distribution copies made available to those personnel most affected by the plan. The plan should be in a document control format in order to maintain a historical record of major revisions in the project quality assurance program.

Quality assurance plan format - The format used for the QA plan will vary from project officer to project officer. The following format is suggested, primarily to highlight points for consideration by the project officer in preparing the plan.

1. Table of Contents. Show in document control and revision format similar to the Handbook table of contents.
2. Project Description. Give a brief project description including references to be consulted for more details.
3. Organization and Responsibility. Give a table or

chart showing key individuals or groups involved in the project quality assurance activities.

4. Project Quality Assurance Program. The following is a minimum that should be contained in QA plans in regard to the elements of quality assurance:

a. Pretest preparation. Include criteria or factors used to select sample collection sites. For ambient air monitoring, a map of the site with the location of air pollution and meteorological sensors should be included; and for source emission monitoring, a flow diagram of the process to be tested and a diagram (with dimensions) of the stack or duct to be tested, plus pretest engineering data on stack velocity, stack temperature, etc.

b. Sample collection. Include a copy of each sensor sample collection procedure. For manual sample collection, include a procedure for sample handling and storage requirements. If standard procedures are used, reference to these procedures will suffice (e.g., EPA Method 5 - Stationary Source Emission Monitoring for Particulates).

c. Sample analysis. Include a copy of all analytical procedures. If standard procedures are used, reference to these procedures will suffice.

d. Data reporting. Include a copy of all forms for reporting data and an example calculation for each pollutant measurement.

e. Data validation. Include a list of criteria that should be used to validate data.

f. Calibration. Include a copy of all calibration procedures. If standard procedures are used, reference to these procedures will suffice. By sensor, provide: (1) a schedule for calibration; (2) description of type of standard; and (3) procedure planned to establish traceability of calibration standards to available standards of higher quality.

g. Audits. Include a plan for performance and system (on-site inspection) auditing. For performance audits, recommend a frequency for an initial audit by sensor.

h. Interlaboratory tests. Any planned or anticipated participation in interlaboratory tests should be listed.

i. Preventive maintenance. Include a copy of preventive maintenance procedures for sensor and auxiliary equipment. Where vendor preventive maintenance procedures are available, reference to these will suffice. A preventive maintenance schedule should be recommended.

j. Other. Any other quality assurance elements, e.g., procurement quality control, configuration control, training, etc., that are applicable to the project should also be included.

5. Forms. Any forms that are not included in part 4a through 4g should be included.

6. Distribution list. A complete list of all organizations and persons to whom the QA plan has been distributed should be included.

APPENDIX A

DEFINITION OF TERMS

A.1 INDEX

Quality Assurance

	Page
Acceptance Sampling	4
Audit	4
Data Validation	4
Performance Audit	5
Quality	5
Quality Assurance	6
Quality Assurance Manual	6
Quality Assurance Plan	6
Quality Control	7
Internal Quality Control	7
External Quality Control	7
Random Samples	7
Representative Sample	8
Sample	8
Statistical Control Chart	9
Stratified Sample	9
System Audit	10

Statistics

Availability	11
Completeness	11
Confidence Coefficient	11
Confidence Interval	11
Confidence Limits	11
Degrees of Freedom	11
Error	12
Maintainability	12
Measures of Central Tendency	12

- Arithmetic Mean 12
- Geometric Mean 12
- Median 13
- Mode 13

Measures of Dispersion or Variability 13

- Range 13
- Standard Deviation 14
- Variance 14
- Coefficient of Variation (Relative Standard Deviation) 15
- Percent Difference 15

Outlier 16

Random Error 16

Relative Error 16

Reliability 17

Statistical Control Chart Limits 17

Systematic Error 17

Test Variability 17

- Accuracy 17
- Bias 18
- Precision 18
- Repeatability 20
- Replicability 20
- Replicates 21
- Reproducibility 21

Tolerance Limits 21

Testing or Measurement

Blank or Sample Blank 22

- Analytical or Reagent Blank 22
- Dynamic Blank (or Field Blank) 22

Calibration 22

- Dynamic Calibration 23
- Static Calibration 23

Collaborative Tests (or Studies) 23

Functional Analysis 23

Proficiency Testing 24

Ruggedness Testing 24

Spiked Sample 24

Standards in Naturally Occurring Matrix 24

- Standard Reference Material (SRM) 25
- Standard Reference Sample (SRS) 25

Standards Depending upon "Purity" or Established Physical or Chemical Constants 26

- Primary Standard 26
- Secondary Standard 26

Standards Based Upon Usage 26

- Calibration Standard 26
- Quality Control Reference Sample (Or Working Standard) 26

Standardization 26

A.1 DEFINITIONS

Quality Assurance

Acceptance Sampling - The procedures by which decisions to accept or reject a sampled lot or population are made based on the results of a sample inspection. In air pollution work, acceptance sampling could be used when checking a sample of filters for certain measurable characteristics such as pH, tensile strength, or collection efficiency to determine acceptance or rejection of a shipment of filters, or when checking the chemical content of a sample of vials of standard solutions from a lot of vials to be used in an interlaboratory test.

Audit - A systematic check to determine the quality of operation of some function or activity. Audits may be of two basic types: (1) performance audits in which quantitative data are independently obtained for comparison with routinely obtained data in an air pollution measurement system, or (2) system audits of a qualitative nature that consist of an on-site review of a laboratory's quality assurance system, and physical facilities for air pollution sampling, calibration, and measurement.

Data Validation - A systematic effort to review data to identify any outliers or errors and thereby cause deletion or flagging of suspect values to assure the validity of the data for the user. This "screening" process may be done by manual and/or computer methods, and may utilize any consistent technique such as sample limits to screen out impossible values or complicated acceptable relationships of the data with other data.

Performance Audit - Planned independent (duplicate) sample checks of actual output made on a random basis to arrive at a quantitative measure of the quality of the output. Performance audits conducted by an auditor are suggested and tabulated in the method activity matrix under the title "Auditing Procedures." Each method in Volume II (Ambient Air Specific Methods) and in Volume III (Stationary Source Specific Methods) has a method activity matrix. These independent checks are made by an auditor subsequent to the routine checks by a field technician or laboratory analyst. One type of performance audit involves the use of a second (or dual) instrument independently maintained, calibrated, and operated at a given monitoring site by an auditor.

Quality - The totality of features and characteristics of a product or service that bear on its ability to satisfy a given purpose. For air pollution measurement systems, the product is air pollution measurement data and the characteristics of major importance are accuracy, precision, and completeness. For air monitoring systems, "completeness," or the amount of valid measurements obtained relative to the amount expected to have been obtained, is usually a very important measure of quality. The relative importance of accuracy, precision, and completeness depends upon the particular purpose of the user.

Quality Assurance - A system for integrating the quality planning, quality assessment, and quality improvement efforts of various groups in an organization to enable operations to meet user requirements at an economical level. In air pollution measurement systems, quality assurance is concerned with all of the activities that have an important effect on the quality of the air pollution measurements as well as the establishment of methods and techniques to measure the quality of the air pollution measurements. The more authoritative usages differentiate between "quality assurance" and "quality control," quality control being "the system of activities to provide a quality product," and quality assurance being "the system of activities to provide assurance that the quality control system is performing adequately."

Quality Assurance Manual - An orderly assembly of management policies, objectives, principles, and general procedures by which an agency or laboratory outlines how it intends to produce quality data.

Quality Assurance Plan - An orderly assembly of detailed and specific procedures by which an agency or laboratory delineates how it produces quality data for a specific project or measurement method. A given agency or laboratory would have only one quality assurance manual, but would have a quality assurance plan for each of its projects or programs (group of projects using the same measurement methods; for example, a laboratory service group might develop a plan by analytical instrument since the service is provided to a number of projects).

Quality Control - The system of activities designed and implemented to provide a quality product.

Internal Quality Control - The routine activities and checks, such as periodic calibrations, duplicate analyses, use of spiked samples, etc., included in normal internal procedures to control the accuracy and precision of a measurement process. (See Quality Control.)

External Quality Control - The activities which are performed on an occasional basis, usually initiated and performed by persons outside of normal routine operations, such as on-site system surveys, independent performance audits, interlaboratory comparisons, etc., to assess the capability and performance of a measurement process.

Random Samples - Samples obtained in such a manner that all items or members of the lot, or population, have an equal chance of being selected in the sample. In air pollution monitoring the population is usually defined in terms of a group of time periods for which measurements are desired. For 24-hour samplers, the population is usually considered as all of the 365 (or 366) 24-hour calendar day periods in a year. For continuous monitors, the population is often considered as all of the hourly average values obtained (or which could have been obtained) during a particular period of time, usually a calendar year. For either 24-hour or continuous monitors, a single air pollution result from a site could be a sample of the conceptually infinite population of values that might have been obtained at the given site for all possible combinations of equipment, materials, personnel, conditions, etc., that could have existed at that site and time.

Representative Sample - A sample taken to represent a lot or population as accurately and precisely as possible. A representative sample may be either a completely random sample or a stratified sample depending upon the objective of the sampling and the conceptual population for a given situation.

Sample - A subset or group of objects or things selected from a larger set, called the "lot," or "population." The objects or things may be physical such as specimens for testing or they may be data values representing physical samples. Unless otherwise specified, all samples are assumed to be randomly selected. Usually, information obtained from the samples is used to provide some indication or inference about the larger set. Samples rather than the population are examined usually for reasons of economy--the entire population under consideration is usually too large or too inaccessible to evaluate. In cases where destructive testing is performed, sampling is a must--otherwise the entire population would be consumed. In many situations, the population is conceptually infinite and therefore impossible to check or measure.

Statistical Control Chart (Also Shewhart Control Chart) - A graphical chart with statistical control limits and plotted values (usually in chronological order) of some measured parameter for a series of samples. Use of the charts provide a visual display of the pattern of the data, enabling the early detection of time trends and shifts in level. For maximum usefulness in control, such charts should be plotted in a timely manner, i.e., as soon as the data are available.

Stratified Sample (Stratified Random Sample) - A sample consisting of various portions that have been obtained from identified subparts or subcategories (strata) of the total lot, or population. Within each category or strata, the samples are taken randomly. The objective of taking stratified samples is to obtain a more representative sample than that which might otherwise be obtained by a completely random sampling. The idea of identifying the subcategories or strata is based on knowledge or suspicion of (or protection against) differences existing among the strata for the characteristics of concern. The identification of the strata is based on knowledge of the structure of the population, which is known or suspected to have different relationships with the characteristic of the population under study. Opinion polls or surveys use stratified sampling to assure proportional representation of the various strata (e.g., geographic location, age group, sex, etc.). Stratified sampling is used in air monitoring to assure representation of different geographical areas, different days of the week, etc.

System Audit - A systematic on-site qualitative review of facilities, equipment, training, procedures, record-keeping, validation, and reporting aspects of a total (quality assurance) system, to arrive at a measure of the capability and ability of the system. Even though each element of the system audit is qualitative in nature, the evaluation of each element and the total may be quantified and scored on some subjective basis.

Statistics

Availability - The fraction or percentage of time that an item is expected to perform satisfactorily (in the reliability sense), taking into account its reliability and its maintainability, or the percentage of "up time" of an item or piece of equipment, as contrasted to its percentage of inoperative or "downtime."

Completeness - The amount of valid data obtained from a measurement system compared to the amount that was expected to be obtained under correct normal operations.

Confidence Coefficient - The chance or probability, usually expressed as a percentage, that a confidence interval has of including the population value. The confidence coefficients usually associated with confidence intervals are 90, 95, and 99 percent. For a given sample size, the width of the confidence interval increases as the confidence coefficient increases.

Confidence Interval - A value interval that has a designated probability (the confidence coefficient) of including some defined parameter of the population.

Confidence Limits - The outer boundaries of a confidence interval.

Degrees of Freedom - The number of independent deviations used in calculating an estimate of the standard deviation. For example, if n analytical results are available, there are n deviations from the average of the n results. Only (n-1) of these deviations are independent; in getting the average, we imposed the restraint that the n deviations must sum to zero.

Error - The difference between an observed or measured value and the best obtainable estimate of its true value.

Maintainability - The probability that an item that has failed (in the reliability sense) can be restored (i.e., repaired or replaced) within a stated period of time.

Measures of Central Tendency - Measures of the tendency of values in a set of data to be centered at some location. Measures of central tendency are the median, the mode, the arithmetic mean, and the geometric mean.

Arithmetic Mean - The most commonly used measure of central tendency, commonly called the "average." Mathematically, it is the sum of all the values of a set divided by the number of values in the set.

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Geometric Mean - Mathematically, the geometric mean can be expressed in two equivalent ways.

1) $$\left[\prod_{i=1}^{n} X_i\right]^{\frac{1}{n}}$$

or in words, the n^{th} root of the product of all values in a set of n values.

2) $$\log^{-1}\left[\frac{\sum_{i=1}^{n} \log X_i}{n}\right]$$

or in words, the antilogarithm of the arithmetic mean of the logarithms of all the values of a set of n values. (Note: the logarithms may be either natural or base 10, or any base for that matter, providing the operations are consistent, i.e., not mixed-base.)

The geometric mean is generally used when the logarithms of a set of values are nearly normally (Gaussian) distributed, such as is the case for much pollution data.

Median - The middle value of a set of data when the set of data are ranked in increasing or decreasing order. If there are an even number of values in the set, the median is the arithmetic average of the two middle values.

Mode - The value or values occurring most frequently in a set or sample of data.

Measures of Dispersion or Variability - Measures of the differences, scatter or variability of values of a set of numbers. Measures of the dispersion or variability are the range, the standard deviation, the variance, and the coefficient of variation.

Range - The difference between the maximum and minimum values of a set of values. When the number of values is small (i.e., 12 or less), the range is a relatively sensitive (efficient) measure of variability.

As the number of values increases above 12, the efficiency of the range (as an estimator of the variability) decreases rapidly. The range or difference between two paired values is of particular importance in air pollution measurements, since in many situations duplicate analyses or measurements are performed as a part of the quality assurance program.

Standard Deviation - The square root of the variance of a set of values:

$$s = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n - 1}}$$

If the values represent a sample from a larger population:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} (X_i - \mu)^2}{N}}$$

where μ is the true arithmetic mean of the population. The property of the standard deviation that makes it most practically meaningful is that it is in the same units as the values of the set, and universal statistical tables for the normal (and other) distributions are expressed as a function of the standard deviation. Mathematically, the tables could just as easily be expressed as a function of the variance.

Variance - Mathematically, for a sample, the sum of squares of the differences between the individual values of a set and the arithmetic mean of the set, divided by one less than the number of values.

$$s^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n - 1}$$

For a population, the sum of squares of deviations from the arithmetic mean, divided by the number of values in the population.

$$\sigma^2 = \frac{\sum_{i=1}^{N} (X_i - \mu)^2}{N} ,$$

where μ is the true arithmetic mean of the population.

Coefficient of Variation (Relative Standard Deviation) - A measure of precision calculated as the standard deviation of a set of values divided by the average. It is usually multiplied by 100 to be expressed as a percentage.

$$\hat{CV} = \frac{s}{\bar{x}} \times 100 \quad \text{for a sample, or}$$

$$CV = \frac{\sigma}{\mu} \times 100 \quad \text{for a population.}$$

Examples of the computations for range, standard deviation, variance and coefficient of variation are presented in Appendix C, pages 11-14.

Percent Difference - If two independent measurements of the same characteristics are available, it is convenient to use the percent difference instead of the coefficient of variation, i.e.

$$\text{Percent Difference} = \frac{|X_1 - X_2|}{(X_1 + X_2)/2} \times 100 = \frac{\text{Range}}{\bar{X}} \times 100$$

This measure of dispersion is readily calculated and is proportional to the coefficient of variation when there are only two measurements of the same characteristic. The percent difference is frequently used when audit data are available based on an independent instrument/operator.

Outlier - An extreme value that questionably belongs to the group of values with which it is associated. If the chance probability of its being a valid member of the group is very small, the questionable value is thereby "detected" and may be eliminated from the group on a statistical basis.

Random Error - Variations of repeated measurements that are random in nature and individually not predictable. The causes of random error are assumed to be indeterminate or nonassignable. The distribution of random errors is generally assumed to be normal (Gaussian).

Relative Error - An error expressed as a percentage of the true value or accepted reference value. All statements of precision or accuracy should indicate clearly whether they are expressed in absolute or relative sense. (This gets complicated when the absolute value is itself a percentage as is the case for many chemical analyses.)

Reliability (General) - The ability of an item or system to perform a required function under stated conditions for a stated period of time. (Specific) - The probability that an item will perform a required function under stated conditions for a stated period of time.

Statistical Control Chart Limits - The limits on control charts that have been derived by statistical analysis and are used as criteria for action, or for judging whether a set of data does or does not indicate lack of control.

Systematic Error - The condition of a consistent deviation of the results of a measurement process from the reference or known level. The cause for the deviation, or bias, may be known or unknown, but is considered "assignable." By assignable is meant that if the cause is unknown, it should be possible to determine the cause. See Bias.

Test Variability

Accuracy - The degree of agreement of a measurement (or an average of measurements of the same thing), X, with an accepted reference or true value, T, usually expressed as the difference between the two values, X - T, or the difference as a percentage of the reference or true value, 100(X-T)/T, and sometimes expressed as a ratio, X/T.

Bias - A systematic (consistent) error in test results. Bias can exist between test results and the true value (absolute bias, or lack of accuracy), or between results from differenct sources (relative bias). For example, if different laboratories analyze a homogeneous and stable blind sample, the relative biases among the laboratories would be measured by the differences existing among the results from the different laboratories. However, if the true value of the blind sample were known, the absolute bias or lack of accuracy from the true value would be known for each laboratory. See Systematic Error.

Precision - A measure of mutual agreement among individual measurements of the same property, usually under prescribed similar conditions. Precision is most desirably expressed in terms of the standard deviation but can be expressed in terms of the variance, range, or other statistic. Various measures of precision exist depending upon the "prescribed similar conditions." (See Replicability, Repeatability, Reproducibility.)

Measures of precision must be qualified or explained in terms of possible sources of variability to make them most meaningful and useful. This is particularly true for repeatability. For example, the following tabulation reflects the requirements of the above definition:

Source of variability	Replicability	Repeatability	Reproducibility
Specimen (subsample)	Same or different	Same or different	Most likely different
Sample	Same	Same	Same
Analyst	Same	At least one of these must be different	Different
Apparatus	Same		Different
Day	Same		Same or different
Laboratory	Same	Same	Different

In the above tabulation, the essential requirement for repeatability is that the same sample must be analyzed by the same laboratory but under different conditions. The situation may be single analyst or multianalyst, single apparatus or multiapparatus, and single day or multiday, or any of the seven possible combinations involving at least one multifactor, each of which would result in different measures of precision. Also, for replicability, repeatability, and reproducibility, the situation may be single specimen or multispecimen, depending usually upon the physical limitations involved. For further detailed discussion, see ASTM Method E177-71, "Use of Terms Precision and Accuracy as Applied to Measurement of a Property of a Material."

Dr. John Mandel, in his paper, "Repeatability and Reproducibility," Materials Research Standards, August 1971, Vol. 11, No. 8, pages 8-16, defines repeatability and reproducibility in the specific sense of an upper confidence limit on expected differences between two test values. In the case of repeatability, the differences are those between two test values at the same laboratory, and in the case of reproducibility, the difference between two test values--one from one laboratory and the second from another laboratory. It is important that the distinction be made between precision measured as a standard deviation and precision expressed as an upper confidence limit of differences between two values since they are frequently used. There is, however, a definite relationship between the two measures. For example, the upper 95 percent confidence limit on differences between two values is 2.77 times the standard deviation.

Repeatability - The precision, usually expressed as a standard deviation, measuring the variability among results of measurements at different times of the same sample at the same laboratory. The unit of time should be specified, since within-day repeatability would be expected to be smaller than between-day repeatability.

Replicability - The precision, usually expressed as a standard deviation, measuring the variability among replicates. It could also be expressed as an upper confidence limit for the difference between two replicates as discussed above.

Replicates - Repeated but independent determinations of the same sample, by the same analyst, at essentially the same time and same conditions. Care should be exercised in considering replicates of a portion of an analysis and replicates of a complete analysis. For example, duplicate titrations of the same digestion are not valid replicate analyses, although they may be valid replicate titrations. Replicates may be performed to any degree, e.g., duplicates, triplicates, etc.

Reproducibility - The precision, usually expressed as a standard deviation, measuring the variability among results of measurements of the same sample at different laboratories.

Tolerance Limits - A particular type of confidence limit used considerably in quality control work where the limits apply to a percentage of the individual values of the population.

Testing or Measurement

Blank or Sample Blank - A sample of a carrying agent (gas, liquid, or solid) that is normally used to selectively capture a material of interest, and that is subjected to the usual analytical or measurement process to establish a zero baseline or background value, which is used to adjust or correct routine analytical results.

Analytical or Reagent Blank - A blank used as a baseline for the analytical portion of a method. For example, a blank consisting of a sample from a batch of absorbing solution used for normal samples, but processed through the analytical system only, and used to adjust or correct routine analytical results.

Dynamic Blank (or Field Blank) - A blank that is prepared, handled, and analyzed in the same manner as normal carrying agents except that it is not exposed to the material to be selectively captured. For example, an absorbing solution that would be placed in a bubbler tube, stoppered, transported to a monitoring site, left at the site for the normal period of sampling, returned to the laboratory, and analyzed.

Calibration - Establishment of a relationship between various calibration standards and the measurements of them obtained by a measurement system, or portions thereof. The levels of the calibration standard should bracket the range of levels for which actual measurements are to be made.

Dynamic Calibration - Calibration of a measurement system by use of calibration material having characteristics similar to the unknown material to be measured. For example, the use of a gas containing sulfur dioxide of known concentrations in an air mixture could be used to calibrate a sulfur dioxide bubbler system.

Static Calibration - The artificial generation of the response curve of an instrument or method by use of appropriate mechanical, optical, electrical, or chemical means. Often a static calibration checks only a portion of a measurement system. For example, a solution containing a known amount of sulfite compound would simulate an absorbing solution through which has been bubbled a gas containing a known amount of sulfur dioxide. Use of the solution would check out the analytical portion of the pararosaniline method, but would not check out the sampling and flow control parts of the bubbler system.

Collaborative Tests (or Studies) - The evaluation of a new analytical method under actual working conditions through the participation of a number of typical or representative laboratories in analyzing portions of carefully prepared homogeneous samples.

Functional Analysis - A mathematical analysis that examines each aspect of the measurement system (sampling and analysis) in order to quantitate the effect of sources of error. Functional analyses have been conducted on each method of this Handbook, Volumes II and III, to identify these parts of the system that should be audited or receive special quality assurance attention. A functional analysis is usually performed prior to a ruggedness test in order to determine those variables which should be studied experimentally.

Proficiency Testing - Special series of planned tests to determine the ability of field technicians or laboratory analysts who normally perform routine analyses. The results may be used for comparison against established criteria, or for relative comparisons among the data from a group of technicians or analysts.

Ruggedness Testing - A special series of tests performed to empirically determine the sensitivity (hopefully, to confirm the insensitivity) of a measurement system to variations of certain factors suspected to affect the measurement system.

Spiked Sample - A normal sample of material (gas, solid, or liquid) to which is added a known amount of some substance of interest. The extent of the spiking is unknown to those analyzing the sample. Spiked samples are used to check on the performance of a routine analysis or the recovery efficiency of a method.

Standards in Naturally Occurring Matrix - Standards relating to the pollutant measurement portions of air pollution measurement systems may be categorized according to matrix, purity, or use. Standards in a naturally occurring matrix include Standard Reference Materials and Standard Reference Samples.

Standard Reference Material (SRM) - A material produced in quantity, of which certain properties have been certified by the National Bureau of Standards (NBS) or other agencies to the extent possible to satisfy its intended use. The material should be in a matrix similar to actual samples to be measured by a measurement system or be used directly in preparing such a matrix. Intended uses include (1) standardization of solutions, (2) calibration of equipment, and (3) monitoring the accuracy and precision of measurement systems.

Standard Reference Sample (SRS) - A carefully prepared material produced from or compared against an SRM (or other equally well characterized material) such that there is little loss of accuracy. The sample should have a matrix similar to actual samples used in the measurement system. These samples are intended for use primarily as reference standards (1) to determine the precision and accuracy of measurement systems, (2) to evaluate calibration standards, and (3) to evaluate quality control reference samples. They may be used "as is" or as a component of a calibration or quality control measurement system. Examples: An NBS certified sulfur dioxide permeation device is an SRM. When used in conjunction with an air dilution device, the resulting gas becomes an SRS. An NBS certified nitric oxide gas is an SRM. When diluted with air, the resulting gas is an SRS.

Standards Depending upon "Purity" or Established Physical or Chemical Constants

Primary Standard - A material having a known property that is stable, that can be accurately measured or derived from established physical or chemical constants, and that is readily reproducible.

Secondary Standard - A material having a property that is calibrated against a primary standard.

Standards Based upon Usage

Calibration Standard - A standard used to quantitate the relationship between the output of a sensor and a property to be measured. Calibration standards should be traceable to Standard Reference Materials or primary standard.

Quality Control Reference Sample (or Working Standard) - A material used to assess the performance of a measurement or portions thereof. It is intended primarily for routine intralaboratory use in maintaining control of accuracy and would be prepared from or traceable to a calibration standard.

Standardization - A physical or mathematical adjustment or correction of a measurement system to make the measurements conform to predetermined values. The adjustments or corrections are usually based on a single-point calibration level.

Bibliography

The following sources were reviewed in the development of Appendix A. The definitions have been prepared with particular meaning and application to air pollution measurement systems, and may not agree in some details with definitions cited in the bibliography. The purpose of Appendix A is to promote a more common understanding and usage of these basic terms for air pollution measurement activities.

1. Akland, G.A., Environmental Protection Agency, Quality Assurance and Environmental Monitoring Laboratory, Research Triangle Park, North Carolina 27711 document prepared for Symposium of Trace Element Analysis, May 1973.

2. The Analyst, Vol. 90, No. 1070, p. 252. Analytical Methods Committee, report prepared by the Analytical Standards Sub-Committee (Sodium Carbonate as a Primary Standard in Acid-Base Titrimetry).

3. Environmental Protection Agency, 999-AP-15 (or 999-WP-15), Environmental Measurements Symposium, Valid Data and Logical Interpretation.

4. Environmental Protection Agency, APTD-0736, Field Operations Guide for Automatic Air Monitoring Equipment.

5. Environmental Protection Agency, APTD-1132, Quality Control Practices in Processing Air Pollution Samples.

6. ASQC Standard A1-1971 (ANSI Standard Z1.6-1971), Definitions, Symbols, Formulas, and Tables for Control Charts.

7. ASQC Standard A2-1971 (ANSI Standard Z1.7-1971), Definitions and Symbols for Acceptance Sampling by Attributes.

8. ASQC Standard A3-1971 (ANSI Standard Z1.7-1971), Glossary of General Terms Used in Quality Control.

9. ASTM, Glossary of ASTM Definitions ASTM Committee E-8 on Nomenclature and Definitions, 2nd Ed., 1973, ASTM, 1916 Race Street, Philadelphia, Pennsylvania 19103.

10. ASTM Standard Recommended Practice E117-71. Use of the Terms Precision and Accuracy as Applied to Measurement of a Property of a Material.

11. ASTM Standard Recommended Practice E180-67. Developing Precision Data on ASTM Methods for Analysis and Testing of Industrial Chemicals.

12. Duncan, A.J., "Quality Control and Industrial Statistics," 3rd Ed., 1965, Richard D. Irwin Inc., Homewood, Illinois.

13. "Glossary of Terms Used in Quality Control," European Organisation for Quality Control, Rotterdam, Netherlands, 1972, 3rd Ed.

14. Federal Register, Vol. 36, No. 84, April 30, 1971, "National Primary and Secondary Ambient Air Quality Standards."

15. Federal Register, "Ambient Air Monitoring Equivalent and Reference Methods," October 12, 1973, Vol. 38, No. 197, Part II.

16. Feigenbaum, A.V., Total Quality Control, Engineering and Management. McGraw-Hill (1961).

17. Department of Health, Education and Welfare, Public Health Service, National Institute for Occupational Safety and Health, Division of Laboratories and Criteria Development, Cincinnati, Ohio 45202.

18. National Institute for Occupational Safety and Health, Cincinnati, Ohio 45202, Industrial Hygiene Service Laboratory Quality Control Manual, Technical Report No. 78 (Draft).

APPENDIX B

NOMENCLATURE

This appendix contains a list of the symbols which are used throughout the Appendices.

a,b - intercept, slope of best fit linear equation by the method of least squares, $Y = a + bX$.

A - factor used in estimating the mean by a confidence interval based on the range, (See Table E.3 for values of A).

A_1 - factor for computing the control chart for $\bar{X}$ given $\bar{s}$ and $\bar{\bar{X}}$, i.e., $UCL_{\bar{X}} = \bar{\bar{X}} + A_1\bar{s}$, $LCL_{\bar{X}} = \bar{\bar{X}} - A_1\bar{s}$.

A_2 - factor used in constructing control chart for $\bar{X}$, given $\bar{\bar{X}}$ and $\bar{R}$, i.e, $UCL_{\bar{X}} = \bar{\bar{X}} + A_2\bar{R}$, $LCL_{\bar{X}} = \bar{\bar{X}} - A_2\bar{R}$.

A - availability (only in Appendix L).

b' - slope of best fit line through the origin.

$\hat{CV}$ - coefficient of variation of the sample = $100\ s/\bar{X}$.

CV(or RSD) - coefficient of variation (or relative standard deviation) of the population = $100\ \sigma/\mu$.

c - acceptance number for a single sample plan when sampling by attributes; i.e., if d is less than or equal to c, the lot is accepted.

D - downtime (only in Appendix L).

D - (signed) difference between two measurements = X_1-X_2.

d - number of defectives observed in a sample of n measurements.

%D - percent difference (relative range) between measurements = $\frac{|X_1-X_2|}{(X_1+X_2)/2}$ x 100, i.e., the range divided by the average.

DF (or df) - degrees of freedom.

d_2 - factor to estimate σ given the mean range, values are given in Table C.2.

D_3, D_4, D_5 - factors used in constructing control chart for R, i.e., $UCL_R = D_4\overline{R}$ and $LCL_R = D_3\overline{R}$, $UWL_R = D_5\overline{R}$.

f_i - frequency of the ith group, cell, or interval.

k - number of samples or sets of data averaged.

L - number of laboratories.

$\log_b X$ - logarithm of X using base b, normally b = 10 or b = e = 2.7183 (natural base).

$\overline{\log X}$ - the mean of the logarithms of the X's in a sample, i.e. $\overline{\log X} = \frac{\Sigma \log X}{n}$.

M - maintainability.

M_R - repair time.

M_D - diagnostic time.

n - number in the sample or number of items on test.

n-1 - number of degrees of freedom associated with estimate s^2 of σ^2 based on a sample of size n.

N - population size, if finite, or lot size in acceptance sampling problems.

n ! - n factorial = n (n-1) (n-2)... 2·1, e.g. 5 ! = 5·4·3·2·1 = 120.

$\binom{n}{r}$, C^n_r (or $n^{C}r$) = the number of combinations of n items taken r at-a-time.

$= \frac{n!}{r!(n-r)!}$, e.g. $C^5_0 = \frac{5!}{0!5!} = 1$, $0! = 1$ by definition, $C^5_2 = \binom{5}{2} = \frac{5!}{2!3!} = 10$.

p - fraction of defects (or defective measurements) in the sample.

p' - fraction of defects in the population of measurements sampled.

P - number of intervals for grouped data.

$P(X)$ - the probability of the event X.

$P\ (a < X < b)$ - the probability that X falls between a and b, $a < b$ (a less than b).

r_{11} - test statistic for the largest value, a possible outlier.

R - range of a data set or sample, = largest value less the smallest value of a set of measurements.

$\overline{R}$ - average range for k samples of the same sample size n, $\overline{R} = \Sigma R/k$.

ΣR - sum of the ranges for k samples.

s^2 - sample variance $= \Sigma(X - \overline{X})^2/(n-1)$.

s - sample standard deviation $= \sqrt{s^2}$.

$\overline{s}$ - average s for k groups of data $= \Sigma s/k$.

s_g - geometric standard deviation, = antilog $\{s (\log X)\}$.

s_D - standard deviation of a set of differences of two paired values.

$s (\log X)$ - the standard deviation of the logarithms of the X's in a sample.

$s_{y|X}$ - standard deviation of the observed response from the fitted line (or curve in general), s is frequently used if it is clearly understood from the context that $s_{y|X}$ is the standard deviation of the discussion.

t - time.

$t_{n-1, \alpha}$ - tabulated t value for specified degrees of freedom (DF = n-1) and for which the fraction α of the absolute values of t exceed $t_{n-1, \alpha}$

T_1 - test statistic for the smallest value, a suspect outlier = $|X_1 - \bar{X}|/s = (\bar{X} - X_1)/s$.

T_n - test statistic for the largest value in a sample, a suspect outlier, $T_n = (X_n - \bar{X})/s$.

U - uptime.

UCL (LCL) - upper (lower) control limit.

UWL (LWL) - upper (lower) warning limit.

UCL_R, $UCL_{\bar{X}}$ - upper control limits for R, $\bar{X}$ respectively, the subscript denotes the variables used in the chart.

w - width of confidence interval (used only in Appendix C).

X_i - ith measurement (also used as the ith smallest measurement of a set of measurements arranged in ascending order, see Appendix F).

$\overline{X}$ - sample mean = $\Sigma X/n = (X_1 + X_2 + \cdots + X_n)/n$.

$\tilde{X}$ - median of a sample

$\overline{X}_g$ - geometric mean of a sample of measurements = antilog $(\overline{\log X})$

X - random variable or measured value.

X' - coded value of X, e.g. $X' = (X-a)/c$, where a and c are selected to simplify the calculation of $\overline{X}$ and s.

X_n - largest value in a sample of size n, see Appendix F.

X_1 - smallest value in a sample of size n, see Appendix F.

$\overline{\overline{X}}$ - grand average = $\Sigma\overline{X}/k$

X - independent or controlled variable such as the concentration of NO_2 (Appendix J).

X_p - predicted value of X for an observed value of y, e.g. an analyzer reading y.

y - dependent variable or response variable.

$\overline{y}$ - mean of the y's for the sample = $\Sigma y/n$

Y - predicted mean response

Z (or u) - standard normal variable = $(X - \mu)/\sigma$ where μ and σ are the mean and standard deviation of the normal distributed variable X.

GREEK NOTATION

Letters of the Greek alphabet are commonly used in statistical texts and literature to denote the parameters of the conceptual population of measurements. These are typically unknown values to be estimated on the basis of a sample of measurements taken from the conceptual population. The estimates are denoted by letters of the English alphabet, e.g.

$\overline{X}$ is an estimate of μ

s is an estimate of σ.

Occasionally the Greek letter with a "hat" is used to denote an estimate, e.g. $\hat{\mu}$ or $\hat{\sigma}$ are estimates of μ and σ respectively. These will be used when considering more than one estimate or an estimate different from the standard one.

α, β (alpha, beta) - parameters (intercept and slope) of true linear relationship between the expected or mean value of y and X, i.e. $Y = \alpha + \beta X + \varepsilon$.

ε (epsilon) - random error of measurement associated with the response variable y.

μ (mu) - mean of the population of measurements

$$= \sum_{i=1}^{N} X_i/N \text{ if a finite population.}$$

μ_g - geometric mean of population = antilog (μ{log X}).

ΠX (pi) - $\prod_{i=1}^{n} X_i = X_1 \cdot X_2 \cdots X_n$, that is, the product of the X_i's of the sample.

σ^2 (sigma squared) - population variance = $\Sigma(X-\mu)^2/N$ if N is finite.

σ - population standard deviation = $\sqrt{\sigma^2}$.

$\sigma_X = \sigma\{X\}$ (sigma sub X) - standard deviation of the variable X; often it is necessary to discuss more than one measure of standard deviation; in this case the variable is denoted by a subscript or in braces { }.

$\sigma_{\overline{X}}$ - standard deviation of the sample mean of n independent measurements = $\sigma/\sqrt{n}$, i.e the population standard deviation divided by $\sqrt{n}$.

σ_g - geometric standard deviation of population = antilog ($\sigma\{\log X\}$).

σ_r^2 - variance among replicates within days, labs.

σ_d^2 - variance among days within labs.

σ_ℓ^2 - variance among laboratories.

$\hat{\sigma}$ (σ "hat") - estimate of σ based on R (range), i.e. $\hat{\sigma} = R/d_2$ or $\overline{R}/d_2$.

χ_i (chi) - mid-value of the ith interval (used only in Appendix C).

$\chi^2_{P,DF}$ (chi-squared) - tabulated value of χ^2 distribution for specified DF and 100 Pth percentile.

APPENDIX C

COMPUTATIONAL EXAMPLES OF DESCRIPTIVE STATISTICS

C.1 INTRODUCTION

Statistical methods dealing with procedures for the collection, analysis and interpretation of data can be grouped into two classes; (1) those used to summarize a body of data to make them more meaningful and (2) those used to make generalizations about a large body of possible data from a small body of available data. More simply, the two classes can be referred to as descriptive statistics and statistics used to make inferences. This appendix is devoted to a discussion of the more frequently used descriptive statistics.

C.2 BASIC CONCEPTS

Data to which statistical methods may be applied may be either measurements made on individual elements or counts of the number of elements that possess specific attributes. The totality of measurements of all individual elements or the count of the number of elements with all possible attributes is referred to as the population (or aggregate). The population may consist of a very large number of elements such as the one-hour concentrations of SO_2 for several years or a small number as for example the number of sources in a county that emit more than 10,000 tons of SO_2 in a year.

A statistical sample is a collection of elements selected in some way from the population. Depending upon the way in which the sample is selected, it may or may not provide data that can be used to make useful inferences about the population.

When a value such as the arithmetic mean is calculated from all possible data for the population, it is referred to as a parameter and identified by the Greek letter μ (mu). The arithmetic mean for a sample is referred to as a statistic and identified by the symbol $\overline{X}$. Similarly, the standard deviation for a sample is s and that for a population is σ (sigma).

C.3 DATA PRESENTATION

C.3.1 Frequency Distribution

Data recorded in the order in which they are received are referred to as raw data. As raw data, it is most difficult to perceive the information contained in a large set of data. A very useful method of presenting a body of data is to group the data into intervals and then count the number of data values that fall within each interval. Table C.1 shows the frequency distribution of 162 measurements of the ambient concentration of suspended particulates made over a period of time at a sampling station.

Table C.1 CONCENTRATION OF SUSPENDED PARTICULATES AT STATION X

Concentration, $\mu g/m^3$	Number of values
25 < 50	3
50 < 75	10
75 < 100	14
100 < 125	24
125 < 150	33
150 < 175	31
175 < 200	18
200 < 225	19
225 < 250	8
250 < 275	2
	162

These same data may be presented graphically as a bar chart or histogram as shown in Figure C.1, where the height of the bar is the frequency of values in the corresponding interval.

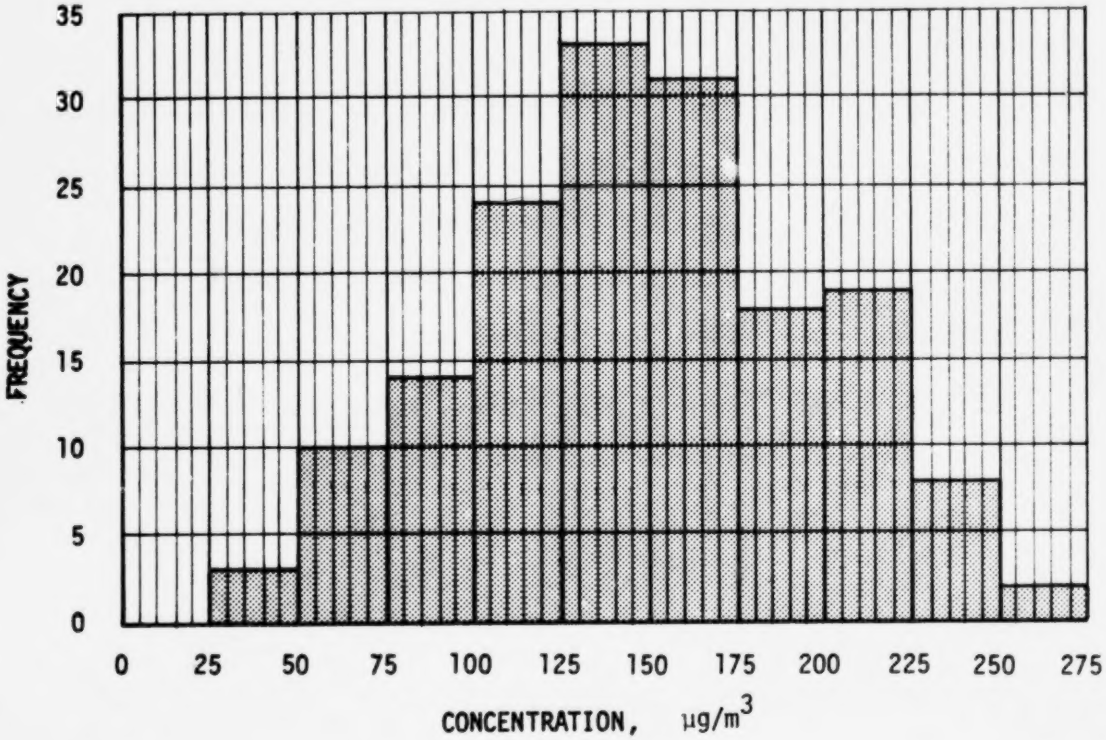


Figure C.1 Frequency distribution of concentrations of TSP.

From Figure C.1, it can now be seen that the ambient concentrations range between 25 and 275 μg/m^3. Also, it can be seen that the most frequently occurring concentrations are in the range of 125-175 μg/m^3.

C.3.2 Measures of Central Tendency

It is often desirable to select a single value to represent a body of data. Such values are referred to as measures of central tendency (or location parameters). Included as measures of central tendency are such parameters as the arithmetic mean, the median, the geometric mean, the mode, the harmonic mean, etc. Several of the more frequently used location parameters are discussed below.

C.3.2.1 Arithmetic Mean - Perhaps the most widely used location parameter is the arithmetic mean. If the frequency distribution for a set of data is nearly symmetrical (as is the case for the data shown in Figure C.1), the arithmetic mean may be the most representative location parameter.

The equations for the arithmetic mean of a finite population and a sample selected from the population are given by,

$$\mu = \frac{1}{N} \sum_{i=1}^{N} X \qquad (1)$$

$$^{*}\overline{X} = \frac{1}{n} \Sigma X \qquad (2)$$

where: μ = population mean

$\overline{X}$ = sample mean

N = number of elements in population

n = number of elements in sample

X = the individual data values

$$\Sigma X = \sum_{i=1} X_i = X_1 + X_2 + \ldots + X_n \qquad (3)$$

* Throughout the text, the notation $\sum_{i=1}^{n} X$ will be replaced by ΣX, with the summation of X for all values in the sample being implied.

Very simply, the arithmetic mean is the sum of the individual values divided by the number of values.

Two examples are presented below to illustrate the computation of the mean for ungrouped data and for grouped data presented as a frequency distribution.

Example C.1 Measurements of the ambient concentration of suspended particulates were made by 12 laboratories as part of a collaborative testing program.[1] Because this is assumed to be a sample, the sample arithmetic mean $\overline{X}$ is calculated

Laboratory	X
1	138
2	125
3	128
4	126
5	127
6	128
7	128
8	108
9	126
10	125
11	125
12	131
ΣX	1515

$$\overline{X} = \frac{1}{12} \Sigma X$$

$$= \frac{1}{12} (1515)$$

$$= 126.2 \ \mu g/m^3$$

Example C.2 The computation of the arithmetic mean for data presented as a frequency distribution is illustrated using the data presented in Ref.1

Concentration	Mid-value χ	No. of values f	$f\chi$
25 < 50	37.5	3	112.5
50 < 75	62.5	10	625.0
75 < 100	87.5	14	1225.0
100 < 125	112.5	24	2700.0
125 < 150	137.5	33	4537.5
150 < 175	162.5	31	5037.5
175 < 200	187.5	18	3375.0
200 < 225	212.5	19	4037.5
225 < 250	237.5	8	1900.0
250 < 275	262.5	2	525.0
		$n = 162$	$\Sigma f \chi = 24075.0$

The equation for the mean for data in a frequency distribution is;

$$\overline{X} = \frac{1}{n} \Sigma f \chi$$

where: f = the number of values in the indicated interval

χ = mid-value of the indicated interval

P = number of intervals

$$\overline{X} = \frac{1}{162} (24075.0)$$

$$\overline{X} = 148.6$$

C.3.2.2 Median - There are situations in which the arithmetic mean may not be the best location parameter to represent a set of data. Consider the following situation in which 6 measurements of the ambient concentration of suspended particulates were 65, 90, 70, 82, 96 and 985 μg/m^3 respectively. The mean of this set of data, 231 μg/m^3, is not truly representative of the typical concentration of suspended particulates. In a situation like this, the median, i.e. the "middle" value may be a more meaningful location parameter.

To determine the median for ungrouped data, it is first necessary to arrange the data in order of magnitude such that $X_1 < X_2 < \ldots < X_n$. When the number of data values is even, the median value is equivalent to the half-way point between the two middle values, i.e., between the $\frac{n}{2}$ and the $\frac{n+2}{2}$ values, and where the number of data values is odd, it is the middle value. With the 6 values above, the median concentration is half-way between the $\frac{6}{2}$ = 3rd and the $\frac{6+2}{2}$ = 4th values.

X_1	65
X_2	70
X_3	82
X_4	90
X_5	96
X_6	985

$$\text{Median} = \frac{82 + 90}{2}$$

$$= 86\ \mu g/m^3$$

For this set of data, 86 $\mu g/m^3$ is a more representative value than 231 $\mu g/m^3$.

C.3.2.3. Geometric Mean - With most air quality measurement data, the frequency distributions tend to be non-symmetrical. The frequency distribution of the 162 values in Example C.4 is "skewed" to the right as shown in Figure C.2.

Under conditions to be discussed in Appendix D in which a frequency distribution is skewed as shown in Figure C.2 the geometric mean may be the most representative location parameter. The geometric mean is defined as the nth root of the product of n values.

where: $\bar{X}_g = (\pi\ X_i)^{1/n}$ (4)

$\bar{X}_g$ = geometric mean

X = the individual data values

and $\pi\ X = X_1 \cdot X_2 \cdots X_n$.

The geometric mean may also be defined as the antilog of the arithmetic mean of the logarithms of the data values;

$$\bar{X}_g = \text{antilog}_b \left[\frac{1}{n} \Sigma \log_b X \right]. \quad (5)$$

Either common logarithms ($\log_{10}$) or natural logarithms ($\log_e$) can be used to calculate the geometric mean.

Example C.3 Calculate the geometric mean for the data presented in Example C.1.

X	$\log_{10}X$
138	2.1398
125	2.0969
128	2.1071
126	2.1003
127	2.1037
128	2.1071
128	2.1071
108	2.0334
126	2.1003
125	2.0969
125	2.0969
131	2.1173
$\Sigma\log_{10}X$ =	25.2075

$$\bar{X}_g = \text{antilog} \left[\frac{1}{n} \Sigma \log_{10}X \right]$$

$$= \text{antilog} \left[\frac{25.2068}{12}\right]$$

$$= \text{antilog}\ [2.10057]$$

$$= 126.1\ \mu g/m^3$$

Example C.4 Calculate the geometric mean for the following data.

Concentration	Mid-value		No. of values	
	X	$\log_{10}X$	f	$f \log_{10}X$
25 < 50	37.5	1.57403	3	4.72209
50 < 75	62.5	1.79588	18	32.32584
75 < 100	87.5	1.94201	37	71.85437
100 < 125	112.5	2.05115	31	63.58565
125 < 150	137.5	2.13830	27	57.73410
150 < 175	162.5	2.21085	14	30.95190
175 < 200	187.5	2.27300	17	38.64100
200 < 225	212.5	2.32736	8	18.61888
225 < 250	237.5	2.37566	5	11.87830
250 < 275	262.5	2.41913	2	4.83826
				335.15039

$$\bar{X}_g = \text{antilog}_{10} \left[\frac{1}{n} \Sigma f \log_{10} X \right]$$

$$= \text{antilog}_{10} \left[\frac{1}{162} (335.15039)\right]$$

$$= \text{antilog}_{10} 2.06883$$

$$= 117.2 \ \mu g/m^3$$

The arithmetic mean for this set of data is 126 $\mu g/m^3$.

C.3.3 Measures of Dispersion

In addition to presenting a location parameter that is representative of a set of data, it is generally important to know the amount of scatter or dispersion of the individual data values. The more widely used measures of dispersion include; the range, the standard deviation, and the variance. In the case of air pollution data the geometric standard deviation is also used.

C.3.3.1 Range - The range is defined as the difference between the largest and the smallest values in a set of data. For Example C.1, the range is 138-108 = 30 $\mu g/m^3$. By definition the range makes use of only two values out of a set of data.

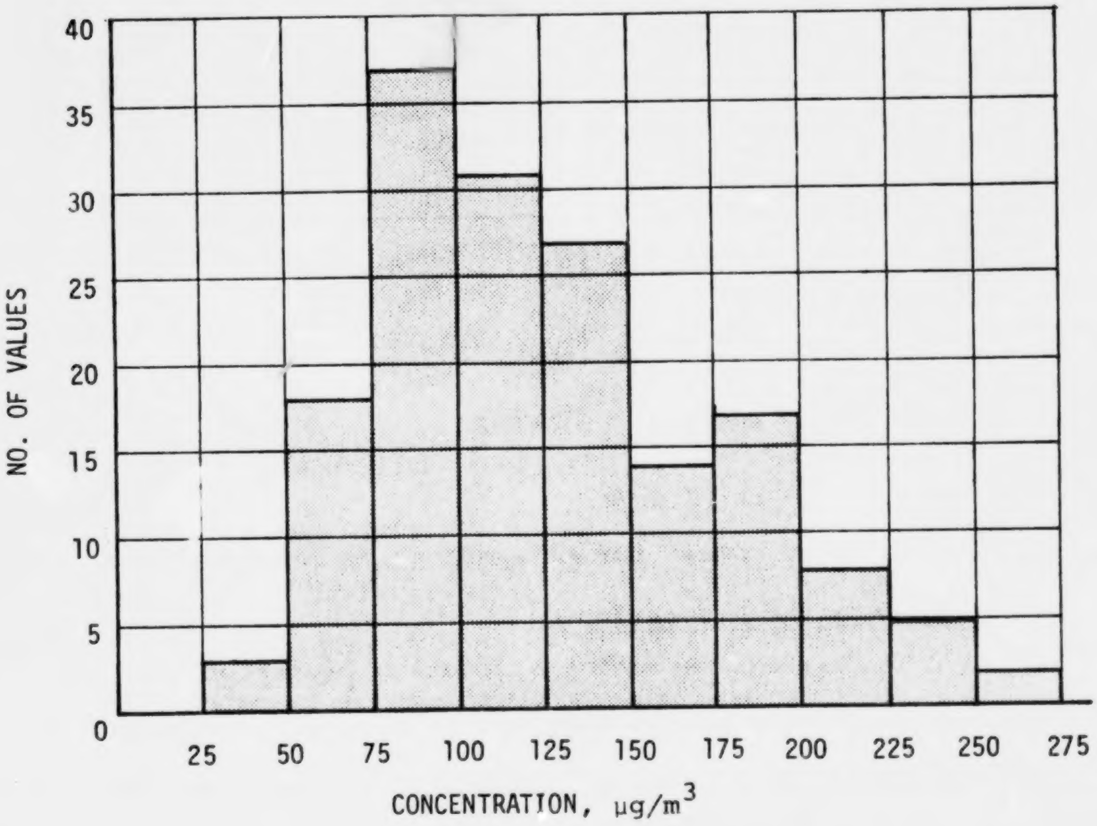


Figure C.2 Frequency distribution of measurements of the concentration of suspended particulates, data for Example C.4.

As such, the range is very sensitive to extreme values. The range has relatively high efficiency compared to the standard deviation when the sample size is small ($2 \leq n \leq 8$). With larger sample sizes, the standard deviation is considerably more efficient than the range and is preferred.

C.3.3.2 Variance and Standard Deviation - The variance of a finite population is defined as the sum of the squares of the deviations of the individual values from the mean divided by the number of values in the population. Denoting the variance by $\sigma^2 = \frac{1}{n} \Sigma (X - \mu)^2$. (6)

The variance of a sample is defined as s^2 and given by

$$s^2 = \frac{1}{n-1} \Sigma (X - \overline{X})^2. \qquad (7)$$

The divisor n-1 is used, rather than n, so that the value of s^2, the sample variance, is an unbiased estimate of σ^2, the population variance.

For computational purposes, with ungrouped data, the equation for s^2 can also be written as

$$s^2 = \frac{\Sigma X^2 - (\Sigma X)^2/n}{n-1} \qquad (8)$$

The equation for s^2 in this form allows one to accumulate the sum (ΣX) and the sums of squares (ΣX^2) very easily on a desk or mini-portable calculator.

Because of the process of squaring, the units of the variance are actually the square of the units of measurement. In order to obtain a statistic with the same units as the original data, the standard deviation, which is defined as the square root of the variance, is used quite frequently.

Example C.5 The variance and standard deviation of the data presented in Example C.1 are calculated below.

Laboratory	X
1	138
2	125
3	128
4	126
5	127
6	128
7	128
8	108
9	126
10	125
11	125
12	131
ΣX	1515
ΣX^2	191,777

$$s^2 = \frac{\Sigma X^2 - \frac{(\Sigma X)^2}{n}}{n-1}$$

$$s^2 = \frac{191{,}777 - \frac{(1515)^2}{12}}{11}$$

$$s^2 = 46.20$$

$$s = 6.8\ \mu g/m^3$$

C.3.3.3 Geometric Standard Deviation - When the data are skewed to the right as illustrated in Figure C.2 and when they are distributed according to the lognormal frequency distribution as described in Appendix D, then the geometric standard deviation s_g is used as a measure of dispersion instead of the standard deviation s. This is defined as

s_g = antilog of the standard deviation of the logarithms of the measurements, i.e.

$$= \text{antilog} \left[\Sigma\ (\log X - \overline{\log X})^2/(n-1) \right]^{1/2}$$

$$= \text{antilog} \left[\frac{\Sigma\ (\log X)^2 - \frac{(\Sigma\ \log X)^2}{n}}{n-1} \right]^{1/2}$$

where the logarithms are taken to any convenient base, preferably the common base 10. Thus the standard deviation of the logarithms of the measurements is computed in the usual manner after transforming each value X (or center value in the case of a frequency distribution such as in Example C.4) to its corresponding log value or log X.

Example C.6 Use the data of Example C.3 to calculate the geometric standard deviation, s_g. For this data,

$$\Sigma\ \log_{10}X = 25.2075$$

$$\Sigma(\log_{10}X)^2 = 52.9579$$

$$s(\log X) = 0.02433$$

$$s_g = \text{antilog } (0.02433) = 1.058$$

$\overline{X}_g$ = 126.07 as obtained in Example C.3.

C.3.3.4 Use of Range to Estimate the Standard Deviation - The range is frequently used to estimate the standard deviation, particularly in control chart activity where the simplicity in calculating the range is desired. Table C.2 gives factors d_2 for dividing the range to estimate the standard deviation. Thus for n = 12, the range would be divided by $d_2 = 3.258$ to estimate σ. For the Example C.1, this estimate would be $30/3.258 = 9.21$, slightly larger than the estimate $s = 6.8$ of σ which is based on all of the data. This relationship between the standard deviation and the range is based on the assumption that the measurements are normally distributed and it can also be used as a quick check on the calculated standard deviation s, i.e. the standard deviation is approximated by

$s \simeq \frac{\text{Range of n measurements}}{d_2}$. It is recommended that a quick rule of thumb be used in all cases to check the calculated s, e.g. one might use

For n between ,	divide the range by d_2 to estimate s
$2 \leq n \leq 5$	$d_2 \simeq 2$
$6 \leq n \leq 15$	$d_2 \simeq 3$
$16 \leq n \leq 50$	$d_2 \simeq 4$
$51 \leq n \leq 200$	$d_2 \simeq 5$
$n > 200$	$d_2 \simeq 6$

Table C.2 FACTORS ASSOCIATED WITH THE RANGE

n	d_2
2	1.128
3	1.693
4	2.059
5	2.326
6	2.534
7	2.704
8	2.847
9	2.970
10	3.078
11	3.173
12	3.258

C.3.3.5 Coefficient of Variation - The coefficient of variation ($\hat{CV}$) as a measure of dispersion is frequently called the relative standard deviation and is computed by:

$$\hat{CV} = \frac{s}{\bar{X}} \times 100 \qquad (9)$$

the ratio of the standard deviation to the mean and multiplied by 100 to convert to a percent of the mean. After some experience with data in a particular field of measurements, typical values of the CV are determined, e.g. 5% or 10% are reasonable values in many of the measurement processes used in measuring pollutant concentrations.

There is one caution that must be kept in mind when stating dispersion in terms of s (absolute terms) or in terms of $s/\overline{X}$ (relative standard deviation). The use of the latter can and does imply that the absolute standard deviation changes with the value $\overline{X}$, whereas stating s implies that it does not change with $\overline{X}$ unless explicitly indicated otherwise. In practice for air pollution measurement the standard deviation does tend to depend on the level of the measurement, but not necessarily with constant proportion over all $\overline{X}$. In many practical problems the CV is essentially constant over the range of interest, and in this case it is the most useful measure of variation. For example, if one were to make replicate analyses of a filter media for lead concentration, the standard deviation of the replicate analyses would tend to increase as the concentration of lead increased and the CV would remain essentially constant.

C.3.3.6 Percent Difference (or Relative Range) - This measure of dispersion is a very useful measure of variation for the special case in which n = 2, i.e. the sample size is two. It is defined by

$$\%D^* = \frac{|X_1 - X_2|}{(X_1 + X_2)/2} \times 100$$

that is, the range of two measurements divided by their average and multiplied by 100 to express the result as a percentage. In the case of two observations the range is directly related to the standard deviation, i.e. $R = \sqrt{2}\, s$, thus the percent difference is $\sqrt{2}$ times the $\hat{CV}$ or relative standard deviation. Because of the ease of computation and the frequency of using repeat or duplicate measurements in air pollution applications, the percent difference is a useful measure of variation. Furthermore, control charts may be applied to this measure when the error of measurement increases with the concentration, say, in preference to an ordinary range chart which assumed that the measurement variation remains constant for the period of application of a control chart and for all levels of concentration.

---

*%D might also be defined for sample sizes other than n=2, e.g. $\%D = R/\overline{X}$ or $\overline{R}/\overline{X}$ (where the latter is used if $n \geq 10$). $|X_1 - X_2|$ is the absolute value of the difference of two measurements.

C.4 COMPUTATIONAL PROCEDURES USING CODED DATA

The computation procedures given in the previous sections of this appendix assume the availability of a simple calculator, preferably one with at least one memory position for computing both the mean and the standard deviation. If the calculations are to be done by hand it is desirable to code the data beforehand and then calculate the mean and standard deviation of these coded data. Then the mean and standard deviation of the original data can be obtained by uncoding, i.e. applying the inverse transformation as required. This approach will be applied to Examples C.1 and C.2 of this appendix to illustrate how the simple computational procedure is applied to both ungrouped and grouped data (frequency distribution).

Example C.7 The data of Example C.1 are arranged in ascending order from the smallest to the largest values as given below. Coded values are denoted by X' and original data by X.

Table C.3 DATA OF EXAMPLE C.1 IN ASCENDING ORDER - COMPUTATION OF MEAN AND VARIANCE USING CODED DATA

Some middle value is arbitrarily selected to coincide with the zero value of the coded value; 126 was chosen for this example.

X	X'=X-126	$(X')^2$
108	-18	324
125	- 1	1
125	- 1	1
125	- 1	1
126	0	0
126	0	0
127	1	1
128	2	4
128	2	4
128	2	4
131	5	25
138	12	144
Sum	3	509

For the coded data X':

$$\bar{X}' = \frac{\Sigma X'}{n} = \frac{3}{12} = 0.25$$

$$s^2_{X'} = \left[\Sigma(X')^2 - (\Sigma X')^2/12\right]/(n-1)$$

$$= \left[509 - (3)^2/12\right]/11$$

$$= 508.25/11 = 46.20$$

$$s_{X'} = \sqrt{46.20} = 6.797$$

For the original data X:

$$\bar{X} = \bar{X}' + 126 = 126.25\ \mu g/m^3$$

$$s_X^2 = s_{X'}^2 = 46.2$$

$$s_X = s_{X'} = 6.797\ \mu g/m^3.$$

Note that the variance of the coded data is the same as that for the uncoded or original data whenever a constant is added to (or subtracted from) each value, without a subsequent division or multiplication.

Example C.8 Using the data of Example C.2 calculate the mean and standard deviation using coded data.

	Mid-value (X)	$X' = \frac{X-162.5}{25}$	Frequency (f)	fX'	$f(X')^2$
	37.5	-5	3	-15	75
	62.5	-4	10	-40	160
	87.5	-3	14	-42	126
Interval →	112.5	-2	24	-48	96
	137.5	-1	33	-33	33
Origin →	162.5	0	31	0	0
	187.5	1	18	18	18
	212.5	2	19	38	76
	237.5	3	8	24	72
	262.5	4	2	8	32
	Sums	---	n=162	$\Sigma fX' = -90$	$\Sigma f(X')^2 = 688$

For the coded data (X'):

$$\overline{X}' = \frac{\Sigma fX'}{n} = -\frac{90}{162} = -0.556$$

$$s^2_{X'} = \left[\Sigma f(X')^2 - (\Sigma fX')^2/n\right]/(n-1)$$

$$= \left[688 - (-90)^2/162\right]/161 = 3.963$$

$$s_{X'} = 1.991$$

For the uncoded or original data (X):

$$\overline{X} = \underset{\substack{\uparrow \\ \text{Interval}}}{25}\ \overline{X}' + \underset{\substack{\uparrow \\ \text{Origin}}}{162.5} = -13.9 + 162.5 = 148.6\ \mu g/m^3$$

$$s^2_X = 25^2 s^2_{X'} = 2476.9$$

$$s_X = \underset{\substack{\uparrow \\ \text{Interval}}}{25} s_{X'} = 49.8\ \mu g/m^3$$

C.5 SUMMARY

In summary, given a set of measurements, they can be summarized by the following quantities or statistics.

Location

Arithmetic Mean $\bar{X} = \frac{\Sigma X}{n}$

$$\bar{X}_g = \text{Antilog}_b \left[\frac{\Sigma \log_b X}{n}\right]$$

$$\text{Median} = \tilde{X} = \begin{cases} \text{Middle value (n odd)} \\ \text{Average of two middle values (n even)} \end{cases}$$

Dispersion

$$s^2 = \frac{\Sigma X^2 - (\Sigma X)^2/n}{n-1}$$

$$s = \sqrt{s^2}$$

$$s_g = \text{antilog} \left[\frac{\Sigma(\log X)^2 - \frac{(\Sigma \log X)^2}{n}}{n-1}\right]^{1/2}$$

R = largest less the smallest value

$\hat{\sigma} = R/d_2$ (See Table C.2)

$$\hat{CV} = \frac{s \times 100}{\bar{X}}$$

$$\%D = \frac{100\,|X_1 - X_2|}{(X_1 + X_2)/2}$$

The use of particular statistics will depend on assumptions concerning the frequency distribution of the measurements as described in the following section. However, the (arithmetic) mean and standard deviation have properties which make them generally useful as measures of the central location and of dispersion of the data and thus as estimates of these same characteristics or parameters of the population.

Additional references [2,3,4] on the subject of descriptive statistics are given at the end of this appendix. Computational forms and instructions are given on pages 23 to 25 for ease of calculation when either (a) one is not familiar with the basic calculations or (b) a reasonably good calculator is not available for use.

C.6 REFERENCES

1. McKee, H. C., Childers, R. E., and Saena, O. S., Jr. Collaborative Study of Reference Method for the Determination of Suspended Particulates in the Atmosphere (High Volume Method), Contract CPA 70-40, SWRI Project 21-2811, June 1971.

2. Bauer, E. L., A Statistical Manual For Chemists, Second Edition, Academic Press, New York, 1971.

3. Guide to Statistical Problem Solving, prepared for EPA by Applied Science Associates, Inc. (A reference for use following the completion of the Basic Environmental Statistics Course.)

4. Daniel, W. W., Biostatistics: A Foundation for Analysis in the Health Sciences, John Wiley and Sons, Inc., New York, 1974.

INSTRUCTIONS FOR CALCULATIONS

1. Enter data values as they occur in the sample unless it is desired to code the data to simplify the calculations. In the latter case rearrange the data from smallest to largest value.

2. For coded data select an origin near the center of the data range, e.g. use the mid-range given by the average of the smallest and largest values rounded off to agree with a specific value.

3. The calculation of the mean is the sum of the observations divided by the number n.

4. The variance, standard deviation calculation is divided into several steps,
   (1) calculate the sum of squares of the values
   (2) calculate the sum of the values, then square this result and divide by n.
   (3) divide the difference, (1)-(2), by n-1, this yields s^2.
   (4) take square root of s^2 to obtain the standard deviation s. (Tables of square roots are available if no calculator is convenient.)

5. The range is the difference, largest value less the smallest value.
   If $n < 10$ calculate one range.
   If $n \geq 10$, divide the data into subgroups where the number in each subgroup is less than 10, e.g. 5 or 6 values, and obtain the range of each subgroup. Then calculate the average range $\bar{R} = \Sigma R/k$, where k is the number of subgroups. $\hat{\sigma} = \bar{R}/d_2$, where where d_2 is obtained in Table C.2 for the sample size n (of each subgroup) is an estimate of the standard deviation σ. In subdividing the data into subgroups use the data in the order in which they were obtained and not in ascending order as for use of coded values. If the total number of observations does not subdivide into equal size subgroups, e.g. if $n = 17$, one observation can be dropped, say first or last measurement, then the 16 observations can be subdivided into 4 groups of 4 each.

6. If a reasonably good calculator is not available the use of coded values can simplify the calculations sufficiently for hand calculation. Two coding approaches are assumed on the form. The first one assumes only a constant shift in the data values, i.e. $X' = X - \text{origin}$, where the origin is selected as suggested in Step (2). In this case, the mean and standard deviation of the coded data are then calculated as in Steps (3) and (4), and for the uncoded data by

   $\bar{X} = \bar{X}' + \text{origin}$, $\quad s_X = s_{X'}$ (no change)

   If the values are further coded by the division (or multiplication) by a constant c (1/c), then

   $\bar{X} = c\bar{X}' + \text{origin}$ and $s_X = cs_{X'}$,

   where c is chosen to simplify the computations.

Single Sample
Worksheet for Calculation of the Mean, Variance, Standard Deviation and Range (Ungrouped Data)

Sample No.	Value* X	Coded* Value X'	Sample No.	Value X	Coded Value X'
1			11		
2			12		
3			13		
4			14		
5			15		
6			16		
7			17		
8			18		
9			19		
10			20		

$\Sigma X =$ ____________

Calculations:

Mean = $\bar{X} = \Sigma X/n =$ ____________

Variance = $s^2 = \Sigma(X-\bar{X})^2/(n-1) = \{\Sigma X^2 - (\Sigma X)^2/n\}/(n-1)$

$\Sigma X^2 =$ __________

$(\Sigma X)^2/n =$ __________

By Subtraction

$\Sigma X^2 - (\Sigma X)^2/n =$ __________

Divide by n-1

$s^2 =$ __________

Standard Deviation = $s = \sqrt{s^2} =$ ____________

Range = R = Largest Value - Smallest Value = ____________

Estimate of σ Based on R is $R/d_2 =$ ______________

(d_2 is taken from Table C.2)

Calculations Using Coded Values:

Use the above steps to calculate $\bar{X}'$ and $s^2_{X'}$, then

$^{**}\bar{X} = \bar{X}' + \text{origin} =$ ______ $^{***}\bar{X} = c\bar{X}' + \text{origin} =$ __________

$s^2_X = s^2_{X'} =$ ______________ $s^2_X = c^2 s^2_{X'} =$ ______________

$s_X = s_{X'} =$ ______________ $s_X = cs_{X'} =$ ______________

*When using coded values, arrange the X values in ascending order.

**It is assumed that X'=X-origin, where the origin is near the center value of the data set.

***It is assumed that X'=(X-origin)/c where c is a constant chosen to simplify the calculation.

Single Sample
Worksheet for Calculation of the Mean, Variance, Standard Deviation and Range (Grouped Data)

Interval No.	Mid-Value X	Coded Value X'	Frequency f	fX'	$f(X')^2$
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
Sums	---	---	n=	$\Sigma fX'$=	$\Sigma f(X')^2$=

Calculations using Coded Values (X'):

Mean $= \overline{X}' = \Sigma fX'/n$

Variance $= s^2_{X'} = \{\Sigma f(X')^2 - (\Sigma fX')^2/n\}/(n-1)$

$\Sigma f(X')^2 =$ ____________

$(\Sigma fX')^2/n =$ ____________

Subtract

Divide by (n-1)

$s^2_{X'} =$ ____________

Standard Deviation $= s_{X'} = \sqrt{s^2_{X'}}$

To calculate the mean and standard deviation for the uncoded data use:

$\overline{X} = c\overline{X}' + \text{origin} =$ ____________________

$s^2_X = c^2 s^2_{X'} =$ ____________

$s_X = cs_{X'} =$ ____________

Where c is the width (or length) of the class interval, i.e. the distance between mid-points assuming all intervals of equal width, and the origin is selected as the mid-point of an interval near the center.

APPENDIX D

PROBABILITY DISTRIBUTIONS

D.1 INTRODUCTION

In Appendix C a frequency distribution was used as a means of presenting a large quantity of data in a meaningful way. If the number of data values become quite large and if the width of the intervals of the frequency distribution is allowed to tend toward zero, it is obvious that the midpoints of the tops of the bar graph will tend to describe a smooth curve. Two major types of continuous frequency distributions are used to describe air pollution data, namely, the normal and log normal frequency distributions. Both of these will be briefly described in this appendix. There are also applications in which the measurement of interest can take on a limited number of distinct values, as for example the number of times during a year when the 3 hour air quality standard for SO_2 was exceeded. In this case the number of such occurrences among n measurements can only be 0, 1, 2, ..., n. The relative frequency of each such occurrence would be an example of a discrete frequency distribution. This class of frequency distribution will not be discussed in this appendix; the reader is referred to other texts on the subject.[1,2]

D.2 NORMAL DISTRIBUTION

The most widely used continuous frequency distribution is the normal (or Gaussian) function. The normal distribution is described by two parameters, the mean (μ) and the standard deviation (σ). Referring to Figure D.1, one can observe that changing the value of σ causes the curve to become more spread out or more peaked. Changing the value of μ merely shifts the curve to the right or left on the horizontal axis.

The area under the normal curve between two specified ordinates can be used to express the probability that a measurement from a normal population would fall in the interval bounded by the two ordinates. Probabilities for selected intervals specified in units of standard deviation are also shown in Figure D.1. Thus, it can be seen that the probability is 0.9544 that a value X selected at random from the normal population will fall in the interval between $\mu - 2\sigma$ and $\mu + 2\sigma$. This statement of probability can also be written in the form

$$P\ (\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0.9544.$$

It should also be obvious from Figure D.1 that

$$P\ (X \geq \mu) = 0.5.$$

Since the normal curve in a particula application depends upon the values of μ and σ, there are an infinite number of possible normal curves. Standard tables of probabilities for the normal curve are constructed for the special case where $\mu = 0$ and $\sigma = 1$. To use such tables it is necessary to rescale

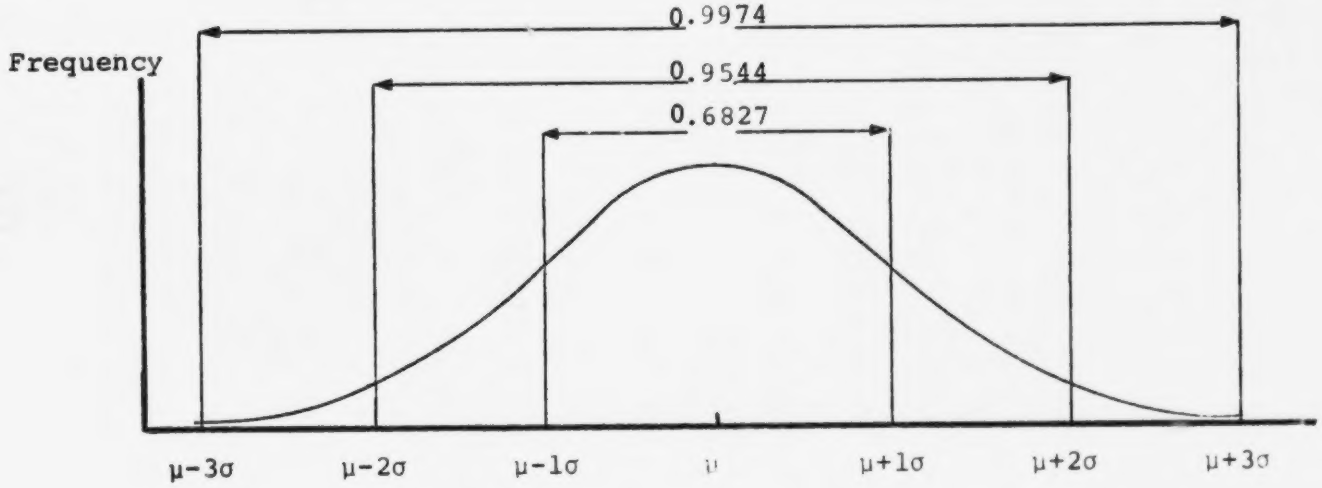
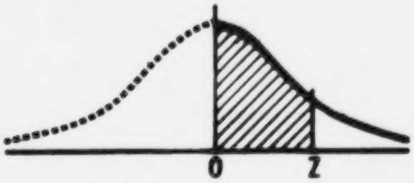

Frequency
0.9974
0.9544
0.6827
μ-3σ
μ-2σ
μ-1σ
μ
μ+1σ
μ+2σ
μ+3σ

Figure D.1 Area under normal curve between specified intervals.

Table D.1[1]



CUMULATIVE NORMAL FREQUENCY DISTRIBUTION

(Area under the standard normal curve from 0 to Z)

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000									

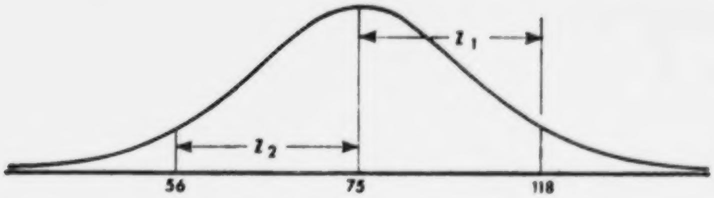
Reproduced with permission from NBS Handbook 91, Experimental Statistics.

the variable of measurement by the following transformation

$$Z = \frac{X - \mu}{\sigma}$$

The quantity Z is usually referred to as a standard normal variate. Probability associated with positive values of Z are presented in Table D.1. This table gives the probability that a value selected at random from the standard normal distribution will fall in the interval $Z = 0$ to $Z = Z_1$.

Example D.1 Suppose the measurement of concentration of a certain pollutant is normally distributed with $\mu = 75$ and $\sigma = 25\ \mu g/m^3$. What is the probability that a measurement made at random will be in the interval between 56 and 118?



$$Z_1 = \frac{X - \mu}{\sigma}$$

$$= \frac{118 - 75}{25}$$

$$= 1.72$$

That is, Z is 1.72 standard deviations larger than the mean value.

$$Z_2 = \frac{56 - 75}{25}$$

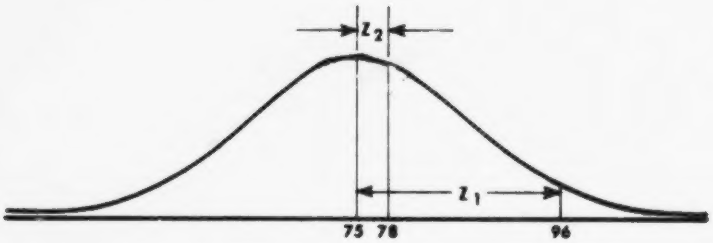
$$= -0.76$$

From Table D.1 the probability for the interval bounded by $Z_1 = 1.72$ is 0.4573 and for $Z_2 = -0.76$ is 0.2764 (the negative sign indicates that the Z_2 lies to the left of the mean). Thus the probability statement for this example is

$$P\ (56 \leq X \leq 118) = 0.4573 + 0.2764$$
$$= 0.7337$$

Thus there is a 73 percent chance that the measured value will fall between 56 and 118 given that the measurements are normally distributed.

Example D.2 For the distribution in Example D.1, what is the probability that a value will lie in the interval between 78 and 96?



$$Z_1 = \frac{96 - 75}{25}$$
$$= 0.84$$

$$Z_2 = \frac{78 - 75}{25}$$
$$= 0.12$$

Referring to Table D.1 to obtain the probabilities for Z_1 and Z_2, the following can be written

$$P\ (78 \le X \le 96) = 0.2995 - 0.0478$$
$$= 0.2517$$

D.3 LOGNORMAL DISTRIBUTION

Frequency distributions of measurements of ambient concentrations of air contaminants have been studied extensively in recent years. Most investigators agree that such distributions are not necessarily normal but they tend to disagree somewhat as to which distribution best describes such data. It has been found that the logarithm of the measured data does tend to be nearly normally distributed. In this situation the distribution of log X is described by the mean and standard deviation of log X. In order to obtain values with units consistent with the measured data, the antilogs are used. Thus, a variable which is lognormally distributed is usually described in terms of the geometric mean (μ_g) and the geometric standard deviation (σ_g) where

$$\mu_g = \text{antilog}\ (\mu\{\log X\}) \qquad (9)$$

$$\sigma_g = \text{antilog}\ (\sigma\{\log X\}) \qquad (10)$$

Example D.3 Suppose that the logarithm of the concentration of total suspended particulates is normally distributed with mean $\mu\{\log X\} = 1.8751(\mu_g=75)$ and the standard deviation $\sigma\{\log X\} = 0.1$. What is the probability that a measurement made at random will fall between 65 and 85 $\mu g/m^3$?

To answer this question, it is first necessary to obtain $\log_{10}65 = 1.8129$ and $\log_{10}85 = 1.9294$. Hence the probability that a measurement falls between 65 and 85 g/m^3 is equal to the probability that the log X falls between 1.8129 and 1.9294, given that the mean $\mu(\log X) = 1.8751$ and $\sigma(\log X) = 0.1$ in log units. Then calculate Z_1 and Z_2 as in examples D.1 and D.2 to obtain

$$Z_1 = \frac{1.9294 - 1.8751}{0.1} = 0.543$$

and

$$Z_2 = \frac{1.8129 - 1.8751}{0.1} = -0.622$$

For these values, the areas for the standard normal curve from Table D.1 are approximately 0.233 (by interpolation) and 0.206. Hence, the probability that a measurement at random falls between 65 and 85 is given by the sum of these two values or about 0.44. See Figure D.2 for a graphical explanation of the above steps.

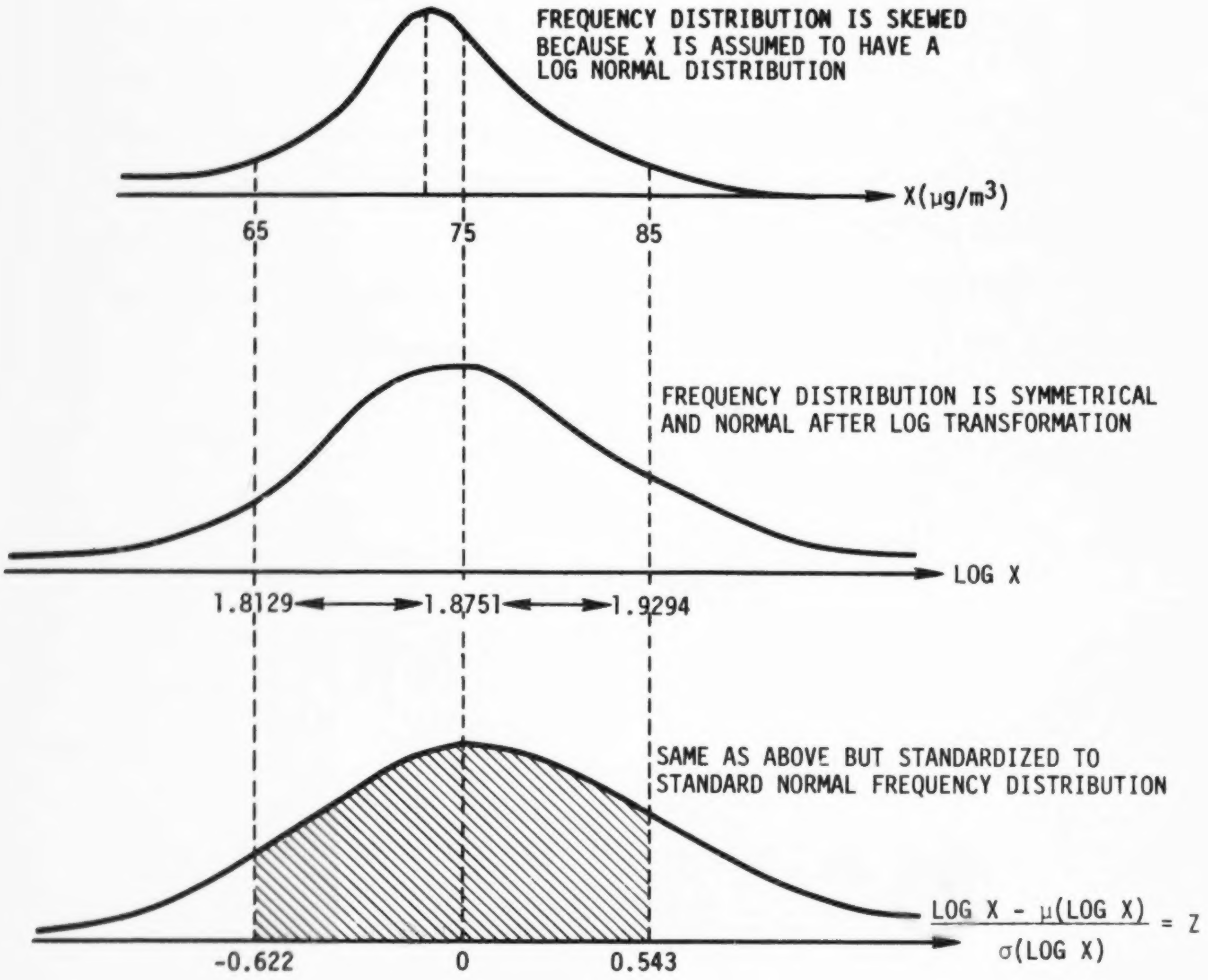

FREQUENCY DISTRIBUTION IS SKEWED BECAUSE X IS ASSUMED TO HAVE A LOG NORMAL DISTRIBUTION
X(μg/m^3)
65
75
85
FREQUENCY DISTRIBUTION IS SYMMETRICAL AND NORMAL AFTER LOG TRANSFORMATION
LOG X
1.8129
1.8751
1.9294
SAME AS ABOVE BUT STANDARDIZED TO STANDARD NORMAL FREQUENCY DISTRIBUTION
$\frac{\text{LOG X} - \mu(\text{LOG X})}{\sigma(\text{LOG X})} = Z$
-0.622
0
0.543

Figure D.2 Illustration of computation of Example D.3.

D.4 DISTRIBUTION OF SAMPLE MEANS

Quite often the process of sampling is used to estimate the mean (μ) of the population. The sample mean ($\overline{X}$) is used as an estimate of the population mean (μ). A major result from statistical theory is that regardless of the shape of the frequency distribution of the original population, the frequency distribution of $\overline{X}$ in repeated samples of size n tends to become normal as n increases. The standard normal distribution can be used to determine probabilities related to the sample mean by the following equation

$$Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Example D.4 Suppose a sample of n = 25 measurements of concentration of air pollutants are made on a population which is normally distributed with mean 60 and standard deviation 15. What is the probability that $\overline{X}$ will lie between 55 and 65?

$$Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$Z_1 = \frac{65 - 60}{\frac{15}{\sqrt{25}}}$$

$$= 1.67$$

$$Z_2 = \frac{55 - 60}{\frac{15}{\sqrt{25}}}$$

$$= -1.67$$

From Table D.1 the probability associated with $Z = 1.67$ is 0.4525. Therefore,

$$P\ (55 \leq \overline{X} \leq 65) = 0.4525 + 0.4525$$

$$= 0.9050$$

Example D.5 Using the population mean, standard deviation and sample size given in Example D.5, what is the probability that $\overline{X}$ would be equal to or greater than 68?

$$Z = \frac{68 - 60}{\frac{15}{\sqrt{25}}}$$

$$= 2.67$$

$$P\ (\overline{X} \geq 68) = 0.5000 - 0.4962$$

$$= 0.0038$$

Example D.6 Suppose that the average of two measurements of a standard reference sample is used as a means of checking on the measurement process by quality control as described in Appendix H. Suppose that the mean and standard deviation of the concentration of the standard reference sample is 75 and 10 μg/m^3 respectively. Between what two values will 95 percent of the averages of samples of size two (n=2) fall?

The standard deviation of the mean of a sample of two is given by $\sigma/\sqrt{n} = 10/\sqrt{2} = 7.1\ \mu g/m^3$. The mean of averages is 75 $\mu g/m^3$; this does not change with sample size. From Table D.1, 47.5 percent of the area under the curve falls between $Z = 0$ and $Z = 1.96$. Hence, the values are determined by

$$\mu \pm Z\ \sigma/\sqrt{n}$$
$$= 75 \pm 1.96\ (10/\sqrt{2})$$

or

$$61.1 \text{ and } 88.9\ \mu g/m^3.$$

This is the process by which quality control limits are determined. For simplicity, they are usually taken to be 2σ or 3σ limits, corresponding to areas of 0.9544 and 0.9974 as given in Figure D.1. In this case the limits are determined by, for example

$$\mu \pm 2(\sigma/\sqrt{n})$$

or

$$\mu \pm 3(\sigma/\sqrt{n}).$$

The 2σ limits are referred to as warning limits and the 3σ as the ordinary control limits. As the sample size n increases the limits become more narrow. In practical applications the sample sizes, $n = 1$ or 2, are most common.

D.5 USE OF PROBABILITY GRAPH PAPER

There are several statistical procedures for checking whether data may be considered to be normally (or lognormally) distributed. One of the most useful procedures in practice is to use normal (or lognormal) probability graph paper, on which the cumulative frequency distribution of a sample of n observations will be approximately a straight line if the data may be considered to be a random sample from the corresponding distribution. An example is given herein to illustrate how to obtain the cumulative frequency distribution and to plot it on the graph paper. The data used are those given in Example C.2. They are repeated here for convenience.

Example D.7 Using the data in Example C.2, obtain the cumulation frequency distribution of the concentration of suspended particulates at station X.

The cumulative frequencies are obtained by cumulating or adding the frequencies in Example C.2 to obtain the total frequency of values of concentration at or below a particular value, in this case the upper value of the corresponding class interval. The relative frequency is then expressed as a percentage of n by dividing the cumulative frequencies by n and multiplying by 100. These are plotted as the abscissae of the points and the ordinates are the corresponding concentrations. In this example, the cumulative frequency distribution follows approximately a straight line and one might assume normality

Concentration X	Cumulative Frequency = Number of Values Less than or Equal to X	Relative Cumulative Frequency (percent)
25	0	0
50	3	1.8
75	13	8.0
100	27	16.7
125	51	31.5
150	84	51.8
175	115	71.0
200	133	82.1
225	152	93.8
250	160	98.8
275	162	100.0

for practical applications. This does not imply that the data are normally distributed, in fact, there is considerable evidence that data on concentration of total suspended particulates tend to follow a lognormal distribution. Lognormal probability paper would be the same as that in Figure D.3 except that the ordinate concentration scale would be in equal steps of the logarithms, i.e. a log scale.

Using the straight line fit to the data (an eye-fitted line in this case), it is possible to estimate the mean and standard deviation of the data in the sample as follows:

(1) An estimate of the mean is obtained by reading the concentration corresponding to the 50th percentile (median value) as shown in the figure, in this case 147 $\mu g/m^3$. Actually $\overline{X}$ = 149 $\mu g/m^3$ from Example C.2.

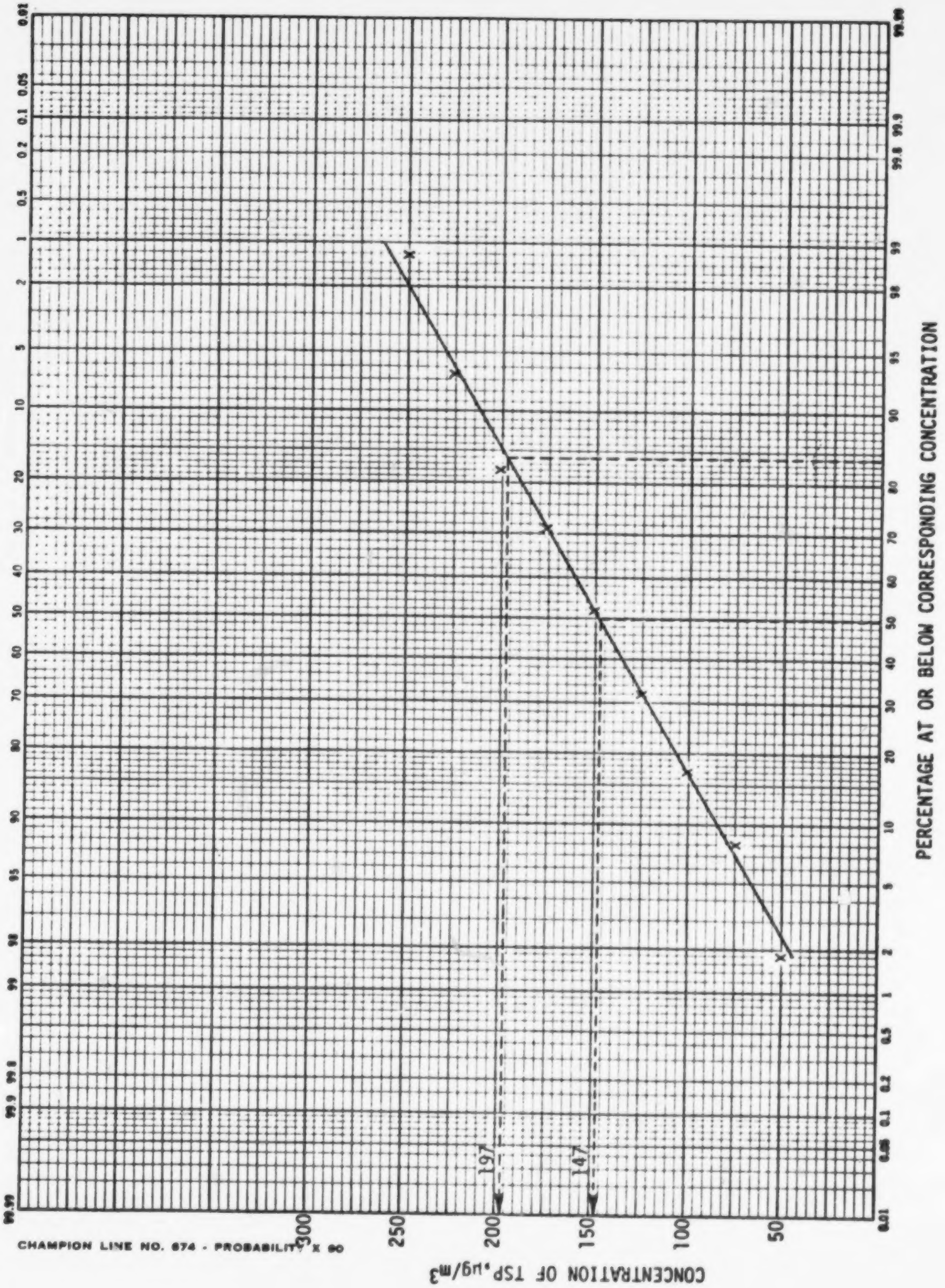


Figure D.3 Cumulative frequency distribution of concentration of TSP ($\mu g/m^3$).

(2) An estimate of the standard deviation is obtained by reading the concentration corresponding to the 84th percentile and subtracting from this the concentration at the mean or 50th percentile, in this case about 197 - 147 or about 50 $\mu g/m^3$ is an estimate of σ. From Table D.1, the area under the standard normal curve between the mean and one standard deviation above the mean, $Z = 1$, is 0.34; or approximately 84 percent of the area lies below the value $\mu + \sigma$ for any normal curve. The calculated standard deviation was obtained in Example C.7 to be 49.8 $\mu g/m^3$. These two sets of estimates compared very closely, in fact, much more closely than one expects in practice. The degree of closeness depends, of course, on how easily one can fit a line to the data as shown in Figure D.3.

D.6 REFERENCES

1. Juran, J.M., Quality Control Handbook, Third Edition, McGraw-Hill Book Co., Inc., New York, 1974.
2. Dixon, W.J., and Massey, F.J., Introduction to Statistical Analysis, McGraw-Hill Book Co., Inc., New York, 1951.

APPENDIX E

ESTIMATION PROCEDURES

E.1 INTRODUCTION

The problems to which statistical methods of analysis are most often applied fall into one of two classes: (1) estimation of one or more unknown parameters for the population from which the sample was selected, and (2) testing hypotheses concerning the population parameters or the validity of the model assumed for the population. Problems of estimation can be further subdivided into those involving point estimates and those involving interval estimation. The problems of testing hypotheses will not be included in this appendix. The reader may refer to texts referenced at the end of this appendix.[1,2,3,4]

E.2 ESTIMATION

E.2.1 Point Estimate

It is often necessary to obtain a single value estimate of a population parameter. For example, the geometric mean ($\bar{X}_g$) of a sample of concentrations of TSP for n equal to 60 days (or every sixth day) is used to estimate the annual geometric mean (μ_g). Similarly, if one is concerned with variability in a set of data, the sample standard deviation (s) is used to estimate the standard deviation (σ) of the population.

Further, in the situation where a linear calibration curve is used to express the relationship between an instrument reading (y) and the concentration of a standard sample (X), the slope (b) and intercept (a) of the fitted line are used as point estimates of the parameters β and α in the model. This procedure is described in Appendix J.

A point estimate is determined from the data for a sample of n observations selected from the population. The procedure for selecting a sample is very critical. The sample must be a representative subset of the total population. Procedures for the selection of a sample have been developed through extensive study[5]. Obviously if the sample is not representative of the population, the point estimate may be a biased estimate of the population parameters. For example, if it is desired to estimate the annual average (or geometric mean) daily concentration of a pollutant, a sample of days must be selected from all days in the year in such a manner as to represent the entire year in terms of daily, weekly and seasonal variations. See Appendix I concerning procedures for selecting a sample.

Quality control procedures are generally based on samples of fewer than 10 observations. In this situation the range, R = MAX. VALUE - MIN. VALUE, is used to derive a point estimate of the population standard deviation. In statistical language, R is said to be nearly as efficient as the sample standard deviation for small samples. The equation for estimating σ from the range is:

$$\hat{\sigma} = \frac{R}{d_2}$$

Values of d_2 for selected sample sizes are presented below:[6]

n	d_2
2	1.128
3	1.693
4	2.059
5	2.326
6	2.534
7	2.704
8	2.847
9	2.970
10	3.078

If the sample size is large (say n greater than or equal to 10), it is desirable to subdivide the sample into k groups of n each, compute an average range $\bar{R}$ and then use the value of d_2 corresponding to the subample size n.

In this case $\hat{\sigma} = \bar{R}/d_2$

E.2.2 Confidence Interval

It should be immediately obvious that the sample means, $\bar{X}_1$ and $\bar{X}_2$, for two independent samples of size n selected at random from a population would most likely not be the same numerical value. Similarly, neither $\bar{X}_1$ nor $\bar{X}_2$ would be expected to be equal to the population mean. In fact, if one were to select a large number of samples of size n, one could construct a frequency distribution of the sample means. The mean of the sample means ($\bar{\bar{X}}$), if the number of independent samples is quite large, would be essentially equal to the population mean μ. The standard deviation of the sample means is given by $\sigma/\sqrt{n}$, where σ is the standard deviation of the population.

Because the sample mean is not likely to be equivalent in value to the population mean, it is common practice to calculate two values A and B such that, there is a given confidence that the unknown value of the population mean will be within the interval. The interval so specified is referred to as a confidence interval. The probability statement for the confidence interval estimate for the population mean (μ) is:

$$P\ (A \leq \mu \leq B) = 1 - \alpha$$

$$P\ (\bar{X} - \frac{st_{n-1,\alpha}}{\sqrt{n}} \leq \mu \leq \bar{X} + \frac{st_{n-1,\alpha}}{\sqrt{n}}) = 1 - \alpha$$

where $\bar{X}$ = sample mean

s = sample standard deviation

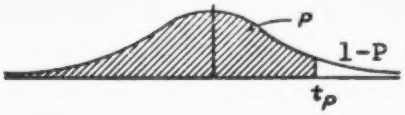
n = sample size

α = risk level (usually 0.10, 0.05, or 0.01)

$t_{n-1,\alpha}$ = value of the Student "t" distribution for n-1 degrees of freedom and risk level α. (See Table E.1)

The risk level (α) is determined on the basis of the consequence which may result from an incorrect decision. Table E.2 is a computational form which can be used in this type of problem.

TABLE E-1 **PERCENTILES OF THE t DISTRIBUTION** *



$1-P = \alpha/2$ (for two-tailed test)

df	$t_{.60}$	$t_{.70}$	$t_{.80}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$
1	.325	.727	1.376	3.078	6.314	12.706	31.821	63.657
2	.289	.617	1.061	1.886	2.920	4.303	6.965	9.925
3	.277	.584	.978	1.638	2.353	3.182	4.541	5.841
4	.271	.569	.941	1.533	2.132	2.776	3.747	4.604
5	.267	.559	.920	1.476	2.015	2.571	3.365	4.032
6	.265	.553	.906	1.440	1.943	2.447	3.143	3.707
7	.263	.549	.896	1.415	1.895	2.365	2.998	3.499
8	.262	.546	.889	1.397	1.860	2.306	2.896	3.355
9	.261	.543	.883	1.383	1.833	2.262	2.821	3.250
10	.260	.542	.879	1.372	1.812	2.228	2.764	3.169
11	.260	.540	.876	1.363	1.796	2.201	2.718	3.106
12	.259	.539	.873	1.356	1.782	2.179	2.681	3.055
13	.259	.538	.870	1.350	1.771	2.160	2.650	3.012
14	.258	.537	.868	1.345	1.761	2.145	2.624	2.977
15	.258	.536	.866	1.341	1.753	2.131	2.602	2.947
16	.258	.535	.865	1.337	1.746	2.120	2.583	2.921
17	.257	.534	.863	1.333	1.740	2.110	2.567	2.898
18	.257	.534	.862	1.330	1.734	2.101	2.552	2.878
19	.257	.533	.861	1.328	1.729	2.093	2.539	2.861
20	.257	.533	.860	1.325	1.725	2.086	2.528	2.845
21	.257	.532	.859	1.323	1.721	2.080	2.518	2.831
22	.256	.532	.858	1.321	1.717	2.074	2.508	2.819
23	.256	.532	.858	1.319	1.714	2.069	2.500	2.807
24	.256	.531	.857	1.318	1.711	2.064	2.492	2.797
25	.256	.531	.856	1.316	1.708	2.060	2.485	2.787
26	.256	.531	.856	1.315	1.706	2.056	2.479	2.779
27	.256	.531	.855	1.314	1.703	2.052	2.473	2.771
28	.256	.530	.855	1.313	1.701	2.048	2.467	2.763
29	.256	.530	.854	1.311	1.699	2.045	2.462	2.756
30	.256	.530	.854	1.310	1.697	2.042	2.457	2.750
40	.255	.529	.851	1.303	1.684	2.021	2.423	2.704
60	.254	.527	.848	1.296	1.671	2.000	2.390	2.660
120	.254	.526	.845	1.289	1.658	1.980	2.358	2.617
∞	.253	.524	.842	1.282	1.645	1.960	2.326	2.576

Adapted by permission of R.A. Fisher and F. Yates, Statistical Tables for Biological, Agricultural and Medical Research, published by Longman Group Ltd., London. (Previously published by Oliver and Boyd, Edinburgh), and by permission of the authors and publishers.

* For two-tailed tests (or symmetrical confidence intervals), use, for example, $t_{0.95}$ for obtaining a 90 percent confidence interval, $t_{0.975}$ for 95 percent confidence, etc.

Table E.2 WORKSHEET FOR CALCULATION OF THE MEAN, VARIANCE, STANDARD DEVIATION, RANGE, AND CONFIDENCE LIMITS

Sample No.	Value X
1	________
2	________
3	________
4	________
5	________
6	________
7	________
8	________
9	________
10	________

Sample No.	Value X
11	________
12	________
13	________
14	________
15	________
16	________
17	________
18	________
19	________
20	________

$\Sigma X =$

Calculations

Mean $= \bar{X} = \Sigma X/n =$ ________

Variance $= s^2 = \Sigma(X-\bar{X})^2/(n-1)$

$= \{\Sigma X^2 - \frac{(\Sigma X)^2}{n}\}/(n-1) =$ ________

$\Sigma X^2 =$ ________

$(\Sigma X)^2/n =$ ________

By Subtraction $\Sigma X^2 - (\Sigma X)^2/n =$ ________

Divide by $n-1$, $s^2 =$ ________

Standard Deviation $= s = \sqrt{s^2} =$ ________

Range = R = Largest Value - Smallest Value = ________

Estimate of σ based on R is R/d_2 = ________

Standard deviation of the mean, $s_{\bar{X}} = s/\sqrt{n}$ = ________

Confidence Limits for Population mean,

$\bar{X} \pm ts_{\bar{X}} =$ ________ $\pm$ ________

Look up the t value in Table E.1 for n-1 degrees of freedom (DF) and desired confidence level, e.g. if n=10, confidence Level is 0.95, (probability of a larger value = 1-0.95=0.05), t = 2.262 for 9 degrees of freedom.

Example E.1 Construct a 95 percent confidence interval estimate of the population mean concentration based upon a sample of 16 observations for which the mean $\overline{X}$ and standard deviation s are 165 and 20 $\mu g/m^3$ respectively.

The value of α is 0.05, and from Table E.1 the value $t_{n-1,\ \alpha}$ is $t_{15,\ 0.05} = 2.131$.

$$\frac{st_{15,\ 0.05}}{\sqrt{n}} = \frac{(20)\ 2.131}{\sqrt{16}} = 11$$

Thus, there is 95 percent confidence that

$$165 - 11 \leq \mu \leq 165 + 11$$

or

$$154 \leq \mu \leq 176\ \mu g/m^3.$$

The interval 154 to 176 $\mu g/m^3$ is defined to be a 95 percent confidence interval for μ.

Example E.2 Using the information for Example E.1, construct a 99 percent confidence interval estimate for μ.

$$t_{15,\ 0.01} = 2.047$$

$$\frac{st_{15,\ 0.01}}{\sqrt{n}} = \frac{(20)\ 2.947}{\sqrt{16}} = 15$$

Thus, there is 99 percent confidence that

$$150 \leq \mu \leq 180\ \mu g/m^3.$$

As the level of confidence increases from 95 to 99 percent, the width of the confidence interval also increases. Similarly, if the level of confidence was to be reduced to 90 percent, the width of the confidence interval would decrease.

The procedures for constructing confidence limits for the standard deviation as well as for regression parameters and other parameters are somewhat more complex than that for the mean. These procedures are presented in many elementary texts on statistical methods.[1,2,3]

The computations for the confidence interval based on the use of the standard deviation are tedious unless a calculator with at least one memory location is available. A simple method requiring only additions and one multiplication in addition to determining the ranges of each of k samples is given in Example E.3 if an easier method is desired.

Example E.3 Suppose that five measurements are made of a standard sample and found to be 44, 50, 47, 50, and 53 $\mu g/m^3$. Assuming no bias in these laboratory measurements, estimate the true mean concentration of the standard with a 95 percent concentration interval using:

(a) the sample mean and standard deviation,

(b) the sample mean and the range.

Solution to (a): The confidence limits are given by

$$\bar{X} \pm \frac{s\, t_{4,0.05}}{\sqrt{n}}$$

where

$$\Sigma X = 244 \qquad \Sigma X^2 = 11,954$$

$$\overline{X} = 48.8\ \mu g/m^3 \qquad \Sigma X^2 - \frac{(\Sigma X)^2}{n} = 46.80$$

$$s = 3.42\ \mu g/m^3 \qquad s^2 = 11.70$$

The 95 percent confidence interval is given by

$$\overline{X} \pm \frac{st_{4,0.05}}{\sqrt{n}}$$

or

$$48.8 \pm \frac{3.42}{\sqrt{5}}(2.776)$$

$$48.8 \pm 4.25.$$

Thus one is 95 percent confident that the true mean falls between 44.5 and 53.1 $\mu g/m^3$.

Solution to (b): In this case the confidence interval is given by $\overline{X} \pm A(\Sigma R)$ where $\overline{X}$ is the mean value, ΣR is the sum of the ranges of k samples each of size n, and A is the factor taken from Table E.3.[2]

For this example $k = 1$, $n = 5$, $A = 0.505$, and $\Sigma R = R = 9$; hence,

$$\overline{X} \pm A(\Sigma R)$$

becomes

$$48.8 \pm 0.505\ (9)$$

or one is 95 percent confident that the true mean falls between 44.2 and 53.4 $\mu g/m^3$.

The solution to (b) is much simpler than (a) in that one only has to perform one addition to obtain the average $\overline{X}$, obtain the range of each sample, and add these ranges to obtain ΣR for the k samples, obtain the appropriate factor A from Table E.3 and apply the formula

$$\overline{X} + A(\Sigma R).$$

This procedure will result in slightly wider intervals *on the average* than the procedure in the solution to (a) above. A computational form is provided in Table E.4 for use in these computations. The only problem that arises in this procedure is that the sample cannot always be subdivided into k groups, each of size n, where n is less than 10. However, in many cases eliminating one or two values (at random or at the end of the data set) will enable this subdivision to be done. In the larger samples this procedure would not reduce the efficiency of the procedure by very much compared to using all of the data.

Example E.4 Suppose that the information is provided later that the true value of the concentration of the standard sample of Example E.3 is 50.1 ± 0.2 $\mu g/m^3$. Are the measurements given in Example E.3 consistent with this information?

It is obvious that they are consistent, but in general, it is only necessary to determine if the confidence interval contains the given or reference value, in this case 50.1 ± 0.2 $\mu g/m^3$. This value falls within the interval given by solutions to either (a) or (b) of Example E.3. In fact it can be inferred with 95% confidence that the measurements are consistent with any reference value between 44.5 and 53.1 $\mu g/m^3$.

E.3 DETERMINATION OF SAMPLE SIZE

In the previous example the width of the confidence interval was given by

$$\frac{2st}{\sqrt{n}}\ n-1,\ 0.05$$

Table E.3

CI = $\bar{X} \pm A(\Sigma R)$. *Confidence Interval for Averages*[*]

Prob.	*k*	*n* = 2	3	4	5	6	7	8	9	10
0.05	1	6.36	1.30	0.719	0.505	0.402	0.336	0.291	0.256	0.232
0.01		31.9	3.00	1.36	0.865	0.673	0.514	0.430	0.379	0.338
0.05	2	0.879	0.316	0.206	0.154	0.125	0.106	0.093	0.084	0.076
0.01		2.11	0.474	0.312	0.227	0.179	0.150	0.131	0.116	0.105
0.05	3	0.360	0.156	0.104	0.079	0.065	0.056	0.049	0.044	0.040
0.01		0.660	0.273	0.150	0.112	0.091	0.077	0.068	0.060	0.054
0.05	4	0.210	0.096	0.065	0.050	0.042	0.036	0.032	0.028	0.026
0.01		0.350	0.142	0.092	0.070	0.057	0.048	0.043	0.038	0.035
0.05	5	0.140	0.066	0.046	0.035	0.030	0.025	0.022	0.020	0.018
0.01		0.226	0.095	0.063	0.049	0.040	0.034	0.030	0.027	0.025
0.05	6	0.102	0.050	0.034	0.027	0.022	0.019	0.017	0.015	0.014
0.01		0.157	0.070	0.047	0.036	0.030	0.026	0.023	0.020	0.019
0.05	7	0.079	0.039	0.027	0.021	0.018	0.015	0.013	0.012	0.011
0.01		0.117	0.055	0.037	0.029	0.024	0.020	0.018	0.016	0.015
0.05	8	0.063	0.032	0.022	0.017	0.014	0.012	0.011	0.010	0.009
0.01		0.094	0.044	0.030	0.023	0.019	0.016	0.014	0.013	0.012
0.05	9	0.053	0.027	0.018	0.014	0.012	0.010	0.009	0.008	0.007
0.01		0.076	0.036	0.025	0.019	0.016	0.014	0.012	0.011	0.010
0.05	10	0.044	0.023	0.016	0.012	0.010	0.009	0.008	0.007	0.006
0.01		0.064	0.031	0.021	0.016	0.014	0.012	0.010	0.009	0.008

[*] Given *k* subgroups of *n* numbers, the confidence interval is $\bar{X} \pm A(\Sigma R)$.

$$A = \frac{t}{kc_1}(\sqrt{kn})$$

Reproduced with permission from Bauer E. L. A Statistical Manual for Chemists, Table 1, page 165, Academic Press, Inc., New York, N. Y. 1971.

Table E.4 CONFIDENCE INTERVAL FOR THE MEAN BASED ON THE RANGE

SAMPLE DATA (k groups - each of size n)

	1	2	3	4	5	6	
1.							
2.							
3.							
4.							
5.							
6.							
7.							
8.							
9.							
10.							
ΣX =							ΣR = ______
R =							$\Sigma\bar{X}$ = ______
$\bar{X}$ =							

Procedure: Subdivide the sample into k groups of n measurements each, where n is less than or equal to 10. It may be necessary to eliminate a measurement to permit the even subdivision into k groups of n each, e.g. if there are 37 measurements, omitting the last one results in six groups of six each, four groups of nine each, or nine groups of four each. When there are alternative possibilities as above, use the subdivision yielding a value nearest to 8 and between 5 and 10 if possible.

1. Calculate the sum, ΣX, range R and mean $\bar{X} = \Sigma X/n$ for each group (column in above table) of n measurements and record above.
2. Compute the sum of the ranges ΣR and means $\Sigma\bar{X}$ over all groups and record above.
3. Calculate the mean over all k groups, $\bar{\bar{X}} = \Sigma\bar{X}/k$ = _______, where k is the number of groups.
4. Obtain factor A from table E.3 for selected probability level, (e.g. if 95 percent confidence is desired and if k = 3, n = 8, then Prob. = 0.05 and A = 0.049). A = _________.
5. Obtain the confidence interval using

   $$\bar{\bar{X}} \pm A(\Sigma R) = \_\_\_\_\_\_\_\_ \pm \_\_\_\_\_\_\_\_ = (\_\_\_\_\_\_\_\_ , \_\_\_\_\_\_\_\_)$$

   and record the lower L and upper limits U in the (___ , ___) above.
6. In conclusion, one infers with 95 (or 99) percent confidence that the mean value for the population sampled falls within the limits L and U, i.e. $L \le \mu \le U$ with the selected confidence level.

when the sample standard deviation s is used and by

$$2A\Sigma R$$

when the range R or ΣR is used. The width of the confidence interval varies from sample to sample according to s and n. Ideally it is desirable to estimate n to yield a confidence which has a practical width, such as a specified percent of the measured value, say 10 percent. In order to determine n precisely, it is necessary to know the standard deviation beforehand. Although this information is not available before the sample is taken, one usually has some previous experience which indicates that the standard deviation is approximately 5 percent of the mean, for example. The computation is given in Example E.5 for estimating the sample size n.

Example E.5 How many measurements should be made of a standard reference sample to obtain the sample mean concentration $\overline{X}$ within 2 percent of the true value μ, assuming σ = 5 percent of the true value, and 95 percent confidence is desired. It is further assumed that there is no measurement bias.

Solution to E.5: The width of the confidence interval is 0.04μ the standard deviation is 0.05μ, where μ is the true mean and thus the width of the interval is given by

$$\frac{2(0.05\mu)}{\sqrt{n}} t_{n-1,0.05} = 0.04\mu$$

or

$$\frac{0.10\ t_{n-1,0.05}}{0.04} = \sqrt{n}$$

or

$$n = 6.25\ t^2_{n-1,0.05}.$$

For n large, $t_{n-1,0.05} \simeq 2$, and hence $n \simeq 25$. Thus if 25 measurements are made of a standard sample and the confidence interval determined, the measured value would fall within about 2 percent of the true value. For example, if $\mu = 50\ \mu g/m^3$, the measured value should fall between 49 and $51 \mu g/m^3$. In general, an approximation to n is given by

$n = 2^2$ (Ratio of the standard deviation $\hat{\sigma}$ to the half-width of the confidence interval both expressed as a percentage of the mean)^2.

i.e

$$n = 4 \left(\frac{\hat{CV}}{W/2}\right)^2$$

where

$\hat{CV} = 100\ \hat{\sigma}/\hat{\mu}$ = estimated coefficient of variation in percent,

$\hat{\sigma}$ = estimated standard deviation

$\hat{\mu}$ = estimated mean level

$W/2$ = half-width of the confidence interval or the maximum percentage deviation desired.

E.4 REFERENCES

1. Youden, W. J., Statistical Methods for Chemists, John Wiley & Sons, Inc., New York, 1951.

2. Bauer, E. L., A Statistical Manual for Chemists, Second Edition, Academic Press, New York, 1971.

3. Mandel, John, The Statistical Analysis of Data, Interscience Publishers, a division of John Wiley & Sons, Inc., New York, 1964.

4. Guide to Statistical Problem Solving, prepared for EPA by Applied Science Associates, Inc. (A reference for use following the completion of the Basic Environmental Statistics Course).

5. Cochran, W. G., Sampling Techniques, John Wiley & Sons, Inc., New York, 1959.

6. Bennett, C. A. and N. L. Franklin, Statistical Analysis On Chemistry and Chemical Industry, John Wiley & Sons, Inc., New York, 1954, (Table 10.1, p. 636).

APPENDIX F

OUTLIERS

F.1 INTRODUCTION

An unusually large (or small) value or measurement in a set of observations is usually referred to as an outlier in the statistical literature. This terminology also has been adopted in the quality control and engineering literature. Some of the reasons for an outlier in data are:

Faulty instrument or component part

Inaccurate reading of record, dial, etc.

Error in transcribing data

Calculation errors

Actual value due to unique circumstances under which the observation(s) was obtained--an extreme manifestation of the random variability inherent in the data.

It is desired to have some statistical procedure to test the presence of an outlier in a set of measurements. The purpose of such tests would be to

(1) Screen data for outliers and hence to identify the need for closer control of the data generating process.

(2) Eliminate outliers prior to analysis of the data. For example, in developing control charts the presence of outliers would lead to limits which are too wide and would make the use of the control charts of minimal, if any, value. In most statistical analysis of data, e.g. regression analysis and analysis of variance, the presence of outliers violate a basic assumption of the analysis.

Incorrect conclusions are likely to result if the outliers are not eliminated prior to analysis. Outliers should be reported, and their omission from analysis should be noted.

(3) Identify the real outliers due to unusual conditions of measurement, e.g. a TSP concentration which is abnormally large due to local environmental conditions during the time of sample collection. Such observations would not be indicative of the average concentration of TSP, and may be eliminated depending on the use of the data. Ideally, these unusual conditions should be recorded on the field data report. Failure to report complete information and unusual circumstances surrounding the collection and analysis of the sample often can be detected by outlier tests.

It will be assumed in the discussion of the appendix that the measurements are normally distributed and that the sample of size n measurements is being studied for the possibility of one or two outliers. If the measurements are lognormally distributed, such as for concentration of TSP, then the logarithm of the data should be taken prior to application of the tests given herein.

F.2 PROCEDURE(S) FOR IDENTIFYING OUTLIERS

Let the set of n measurements be arranged in ascending order and denoted by

$$X_1, X_2, \ldots, X_n$$

where X_i denotes the ith smallest measurement. Suppose that X_n is suspected of being too large, and that a statistical test is to be applied to the particular measurement to determine whether X_n is consistent with the remaining data in the sense that it is reasonable that it is part of the same population of measurements from which the sample is taken. Consider the following data comparing the results of measurements on NO_3 analysis of 10 high-volume filters by two different laboratories. The differences are given and the largest value 0.2 is questionable and is to be tested as a possible outlier.

Example F.1

NO_3 Analysis (mg/filter)

Lab 1	Lab 2	Difference
1.7	2.05	-0.35
2.2	2.4	-0.2
3.9	3.7	0.2
3.3	3.6	-0.3
2.7	3.3	-0.6
3.5	3.8	-0.3
0.9	1.46	-0.56
1.3	1.5	-0.2
6.1	6.4	-0.3
2.9	3.2	-0.3
		-2.91

One test procedure is to compare the value 0.2 with the remaining data through use of a test by Dixon.[1]

$$r_{11} = \frac{X_n - X_{n-1}}{X_n - X_2} = \frac{0.2 - (-0.2)}{0.2 - (-0.56)} = \frac{0.4}{0.76} = 0.53.$$

This value is compared to tabulated values given by Dixon[1], and it is indicated that the value r_{11} is significant at the 0.05 level, i.e. it is larger than 0.477. See Table F.1 for criteria for testing for outliers. Thus one infers that the measured difference is subject to question, and that either Lab 1 obtained an unusually large result, or Lab 2 obtained

Table F.1 DIXON CRITERIA FOR TESTING OF EXTREME OBSERVATION (SINGLE SAMPLE)*

n	Criterion		Significance Level		
			10%	5%	1%
3			.886	.941	.988
4	$r_{10} = \frac{x_2 - x_1}{x_n - x_1}$	if smallest value	.679	.765	.889
5		is suspected;	.557	.642	.780
6		if largest value	.482	.560	.698
7	$= \frac{x_n - x_{n-1}}{x_n - x_1}$	is suspected.	.434	.507	.736
8			.479	.554	.683
9	$r_{11} = \frac{x_2 - x_1}{x_{n-1} - x_1}$	if smallest value	.441	.512	.635
10		is suspected;	.409	.477	.597
	$\frac{x_n - x_{n-1}}{x_n - x_2}$	if largest value is suspected.			
11		if smallest value	.517	.576	.679
12	$r_{21} = \frac{x_3 - x_1}{x_{n-1} - x_1}$	is suspected;	.490	.546	.642
13			.467	.521	.615
	$= \frac{x_n - x_{n-2}}{x_n - x_2}$	if largest value is suspected.			
14		if smallest value	.492	.546	.641
15	$r_{22} = \frac{x_3 - x_1}{x_{n-2} - x_1}$	is suspected.	.472	.525	.616
16			.454	.507	.595
17		if largest value	.438	.490	.577
18	$= \frac{x_n - x_{n-2}}{x_n - x_3}$	is suspected;	.424	.475	.561
19			.412	.462	.547
20			.401	.450	.535
21			.391	.440	.524
22			.382	.430	.514
23			.374	.421	.505
24			.367	.413	.497
25			.360	.406	.489

* Reproduced with permission from W. J. Dixon, "Processing Data for Outliners", Biometrics, March 1953, Vol. 9, No. 1 Appendix, Page 89. (Reference [1])

$X_1 \leq X_2 \leq X_3 \leq \cdots \leq X_{n-2} \leq X_{n-1} \leq X_n$

Criterion r_{10} applies for $3 \leq n \leq 7$

Criterion r_{11} applies for $8 \leq n \leq 10$

Criterion r_{21} applies for $11 \leq n \leq 13$

Criterion r_{22} applies for $14 \leq n \leq 25$

a small value. In any case some investigation of the measurements should be considered, particularly if there are recurring results of this nature which may indicate poor quality data.

Another commonly used test procedure[2], requires additional computation and is given by

$$T_n = (X_n - \bar{X})/s$$

where: $\bar{X}$ is the sample mean

s is the sample standard deviation, i.e.

$s = \{\Sigma(X-\bar{X})^2/(n-1)\}^{1/2}$.

For the data set previously given,

$$\bar{X} = -0.291$$

$$s = 0.22$$

and hence:

$$T_n = \frac{0.2 - (-0.291)}{0.22} = \frac{0.491}{0.22} = 2.23,$$

which is significant at the 0.05 level, i.e. larger than 2.176 which is the tabulated value for this level from Table F.2. This test result is in agreement with the previous one, indicating that this measured difference is larger than expected in sampling from a population which is normally distributed. To aid in computing T_n (or T_1 for testing the smallest value) Table F.3 may be used. In many examples it will be obvious that a particular value is an outlier, whereas in Example F.1 this is not the case. A plot of the data on a straight line, X-axis is often helpful in examining a set of data.

After rejecting one outlier using either T_n or T_1 the analyst may be faced with the problem of considering a

Table F.2 TABLE OF CRITICAL VALUES FOR T (ONE-SIDED TEST OF T_1 OR T_n) WHEN THE STANDARD DEVIATION IS CALCULATED FROM THE SAME SAMPLE

Number of Observations n	Upper .1% Significance Level	Upper .5% Significance Level	Upper 1% Significance Level	Upper 2.5% Significance Level	Upper 5% Significance Level	Upper 10% Significance Level
3	1.155	1.155	1.155	1.155	1.153	1.148
4	1.499	1.496	1.492	1.481	1.463	1.425
5	1.780	1.764	1.749	1.715	1.672	1.602
6	2.011	1.973	1.944	1.887	1.822	1.729
7	2.201	2.139	2.097	2.020	1.938	1.828
8	2.358	2.274	2.221	2.126	2.032	1.909
9	2.492	2.387	2.323	2.215	2.110	1.977
10	2.606	2.482	2.410	2.290	2.176	2.036
11	2.705	2.564	2.485	2.355	2.234	2.088
12	2.791	2.636	2.550	2.412	2.285	2.134
13	2.867	2.699	2.607	2.462	2.331	2.175
14	2.935	2.755	2.659	2.507	2.371	2.213
15	2.997	2.806	2.705	2.549	2.409	2.247
16	3.052	2.852	2.747	2.585	2.443	2.279
17	3.103	2.894	2.785	2.620	2.475	2.309
18	3.149	2.932	2.821	2.651	2.504	2.335
19	3.191	2.968	2.854	2.681	2.532	2.361
20	3.230	3.001	2.884	2.709	2.557	2.385
21	3.266	3.031	2.912	2.733	2.580	2.408
22	3.300	3.060	2.939	2.758	2.603	2.429
23	3.332	3.087	2.963	2.781	2.624	2.448
24	3.362	3.112	2.987	2.802	2.644	2.467
25	3.389	3.135	3.009	2.822	2.663	2.486
26	3.415	3.157	3.029	2.841	2.681	2.502
27	3.440	3.178	3.049	2.859	2.698	2.519
28	3.464	3.199	3.068	2.876	2.714	2.534
29	3.486	3.218	3.085	2.893	2.730	2.549
30	3.507	3.236	3.103	2.908	2.745	2.563
31	3.528	3.253	3.119	2.924	2.759	2.577
32	3.546	3.270	3.135	2.938	2.773	2.591
33	3.565	3.286	3.150	2.952	2.786	2.604
34	3.582	3.301	3.164	2.965	2.799	2.616
35	3.599	3.316	3.178	2.979	2.811	2.628
36	3.616	3.330	3.191	2.991	2.823	2.639
37	3.631	3.343	3.204	3.003	2.835	2.650
38	3.646	3.356	3.216	3.014	2.846	2.661
39	3.660	3.369	3.228	3.025	2.857	2.671
40	3.673	3.381	3.240	3.036	2.866	2.682
41	3.687	3.393	3.251	3.046	2.877	2.692
42	3.700	3.404	3.261	3.057	2.887	2.700
43	3.712	3.415	3.271	3.067	2.896	2.710
44	3.724	3.425	3.282	3.075	2.905	2.719
45	3.736	3.435	3.292	3.085	2.914	2.727
46	3.747	3.445	3.302	3.094	2.923	2.736
47	3.757	3.455	3.310	3.103	2.931	2.744
48	3.768	3.464	3.319	3.111	2.940	2.753
49	3.779	3.474	3.329	3.120	2.948	2.760
50	3.789	3.483	3.336	3.128	2.956	2.768

Reproduced with permission from American Statistical Association.

Use $T_1 = \frac{\overline{X} - X_1}{s}$ when testing the smallest value, X_1.

Use $T_n = \frac{X_n - \overline{X}}{s}$ when testing the largest value, X_n in a sample of n observations. Unless one has prior information about largest values (or smallest values) the risk levels should be multiplied by two for application of the test.

Table F.2(Continued). TABLE OF CRITICAL VALUES FOR T (ONE-SIDED TEST OF T_1 OR T_n) WHEN THE STANDARD DEVIATION IS CALCULATED FROM THE SAME SAMPLE

Number of Observations n	Upper .1% Significance Level	Upper .5% Significance Level	Upper 1% Significance Level	Upper 2.5% Significance Level	Upper 5% Significance Level	Upper 10% Significance Level
51	3.798	3.491	3.345	3.136	2.964	2.775
52	3.808	3.500	3.353	3.143	2.971	2.783
53	3.816	3.507	3.361	3.151	2.978	2.790
54	3.825	3.516	3.368	3.158	2.986	2.798
55	3.834	3.524	3.376	3.166	2.992	2.804
56	3.842	3.531	3.383	3.172	3.000	2.811
57	3.851	3.539	3.391	3.180	3.006	2.818
58	3.858	3.546	3.397	3.186	3.013	2.824
59	3.867	3.553	3.405	3.193	3.019	2.831
60	3.874	3.560	3.411	3.199	3.025	2.837
61	3.882	3.566	3.418	3.205	3.032	2.842
62	3.889	3.573	3.424	3.212	3.037	2.849
63	3.896	3.579	3.430	3.218	3.044	2.854
64	3.903	3.586	3.437	3.224	3.049	2.860
65	3.910	3.592	3.442	3.230	3.055	2.866
66	3.917	3.598	3.449	3.235	3.061	2.871
67	3.923	3.605	3.454	3.241	3.066	2.877
68	3.930	3.610	3.460	3.246	3.071	2.883
69	3.936	3.617	3.466	3.252	3.076	2.888
70	3.942	3.622	3.471	3.257	3.082	2.893
71	3.948	3.627	3.476	3.262	3.087	2.897
72	3.954	3.633	3.482	3.267	3.092	2.903
73	3.960	3.638	3.487	3.272	3.098	2.908
74	3.965	3.643	3.492	3.278	3.102	2.912
75	3.971	3.648	3.496	3.282	3.107	2.917
76	3.977	3.654	3.502	3.287	3.111	2.922
77	3.982	3.658	3.507	3.291	3.117	2.927
78	3.987	3.663	3.511	3.297	3.121	2.931
79	3.992	3.669	3.516	3.301	3.125	2.935
80	3.998	3.673	3.521	3.305	3.130	2.940
81	4.002	3.677	3.525	3.309	3.134	2.945
82	4.007	3.682	3.529	3.315	3.139	2.949
83	4.012	3.687	3.534	3.319	3.143	2.953
84	4.017	3.691	3.539	3.323	3.147	2.957
85	4.021	3.695	3.543	3.327	3.151	2.961
86	4.026	3.699	3.547	3.331	3.155	2.966
87	4.031	3.704	3.551	3.335	3.160	2.970
88	4.035	3.708	3.555	3.339	3.163	2.973
89	4.039	3.712	3.559	3.343	3.167	2.977
90	4.044	3.716	3.563	3.347	3.171	2.981
91	4.049	3.720	3.567	3.350	3.174	2.984
92	4.053	3.725	3.570	3.355	3.179	2.989
93	4.057	3.728	3.575	3.358	3.182	2.993
94	4.060	3.732	3.579	3.362	3.186	2.996
95	4.064	3.736	3.582	3.365	3.189	3.000
96	4.069	3.739	3.586	3.369	3.193	3.003
97	4.073	3.744	3.589	3.372	3.196	3.006
98	4.076	3.747	3.593	3.377	3.201	3.011
99	4.080	3.750	3.597	3.380	3.204	3.014
100	4.084	3.754	3.600	3.383	3.207	3.017

Reproduced with permission from American Statistical Association.

Source: Grubbs, F. E., and Beck, G., Extension of Sample Sizes and Percentage Points for Significance Tests of Outlying Observations, Technometrics, Vol. 14, No. 4, Nov. 1972, pp. 847-854.

Table F.3 CHECK FOR OUTLIERS USING T_1 OR T_n
SINGLE SAMPLE

Sample No.	Value X	Rank*
1	______	______
2	______	______
3	______	______
4	______	______
5	______	______
6	______	______
7	______	______
8	______	______
9	______	______
10	______	______

Sample No.	Value X	Rank*
11	______	______
12	______	______
13	______	______
14	______	______
15	______	______
16	______	______
17	______	______
18	______	______
19	______	______
20	______	______

TOTAL = ΣX = ________________

Calculations:

Mean = $\overline{X} = \Sigma X/n$ = ___________

Variance = $s^2 = \Sigma(X-\overline{X})^2/(n-1) = [\Sigma X^2 - \frac{(\Sigma X)^2}{n}]/(n-1)$ = ________

ΣX^2 = __________

$(\Sigma X)^2/n$ = __________

Difference
$n-1$ = __________

Divide
s^2 = __________

Standard Deviation = $s = \sqrt{s^2}$ = ______________

If Largest Value is Suspect

Calculate

$$T_n = \frac{X_n - \overline{X}}{s} = \_\_\_\_$$

=

If Smallest Value is suspect

Calculate

$$T_1 = \frac{\overline{X} - X_1}{s} = \_\_\_\_$$

=

Compare computed value with tabulated value in Table F.2 for acceptable risk. If $T_n >$ tabulated critical value, the outlier is suspect.

* Rank values from smallest to largest value, e.g. 1 to 20. Actually need to rank three smallest and three largest values.

second outlier. In this case the mean and standard deviation may be recomputed and either T_n or T_1 applied to the sample of n-1 measurements. The user should be aware that the test T_n or T_1 is not theoretically based on repeated use.

Grubbs[2] also gives a test procedure for simultaneously testing the two largest or two smallest values. This procedure is not given here. Grubbs also gives tables for the critical values for this procedure.

The use of the procedures given in Table F.1 requires very little computation and would be preferable on a routine basis. Grubbs[3] gives a tutorial discussion of outliers and is a very good reference to the subject. Another reference[4] on outliers contains Chauvenet's criterion and the appropriate table of critical values. This test is similar to the T_n or T_1 test given here. Chauvenet's criterion for n larger than five rejects outliers more frequently than using the T_n or T_1 test procedure, i.e. it would be less conservative for these larger sample sizes.

The determination of outliers when, for example, several laboratories measure two or more samples in an interlab study, is made by more complex methods.[5,6]

F.3 REFERENCES

1. Dixon, W.J., Processing Data for Outliers, Biometrics, Vol. 9, No. 1, March 1953, pp. 74-89.

2. Grubbs, F.E. and Beck, G., Extension of Sample Sizes and Percentage Points for Significance Tests of Outlying Observations, Technometrics, Vol. 14, No. 4, Nov. 1972, pp. 847-854.

3. Grubbs, F.E., Procedures for Detecting Outlying Observations in Samples, Technometrics, Vol. 11, No. 1, Feb. 1969, pp. 1-21.

4. Guide to Statistical Problem Solving, prepared for U.S. EPA Air Pollution Training Institute.

5. Anscombe, F.J. and Tukey, J.W. The Examination and Analysis of Residuals, Technometrics, Vol. 5, No. 2, May 1963, pp. 141-160.

6. Stefansky, W., Rejecting Outliers in Factorial Design, Technometrics, Vol. 14, 1972, pp. 469-480.

APPENDIX G

TREATMENT OF AUDIT DATA

G.1 AUDIT DATA

One means of checking on the performance of a measurement system or process is to conduct an independent audit of a pertinent portion of the system or of the entire system if possible. In conducting an audit, there will result a set of data collected by the standard test method and a second set of data collected by an audit procedure. The latter may be performed, for example, by an independent operator using the same or different measuring instruments. It is desirable that the two sets of measurements be made, in so far as possible, independently of one another. However, the audit must measure the same characteristic as the standard test measurement. One example of an audit would be to set up a duplicate sampler adjacent to the standard test sampler (but not too close to disturb airflow characteristics), and have different operator(s) perform the collection of the sample and analyze the resultant sample for concentration of TSP in $\mu g/m^3$. Another example would be to have independent weighings of filters made by another analyst or another laboratory.

An audit is usually performed on a sampling basis, that is, checking for example, every tenth filter or one sampled at random from each set of ten. A rate of one out of about fourteen was suggested in the guideline documents.[1] The audit data are then used to infer whether the measurement process is consistent with suggested standards or limits based on available data concerning the expected variation in results of the particular analysis. Many standards were provided with little available data and some were simply engineering/statistical judgments.[1] Thus it is important to analyze audit data not only for determining whether a measurement (process) satisfies suggested standards but also to check the adequacy of the standards. This appendix will treat some of these uses of the audit data using real data.

G.2 ANALYSIS OF AUDIT DATA

Consider the set of data given in Example F.1 of the last appendix. It is repeated here for convenience.

NO_3 Analysis (mg/filter)

No.	Lab 1 (test data)	Lab 2 (audit data)	Difference (D)
1	1.7	2.05	-0.35
2	2.2	2.4	-0.2
3	3.9	3.7	0.2 (Possible outlier)
4	3.3	3.6	-0.3
5	2.7	3.3	-0.6
6	3.5	3.8	-0.3
7	0.9	1.46	-0.56
8	1.3	1.5	-0.2
9	6.1	6.4	-0.3
10	2.9	3.2	-0.3

In Appendix F the difference 0.2 was tested as a possible outlier and determined to be larger than expected by chance with 95 percent confidence. In this section these data are to be analyzed for a different purpose, that is, to determine if the two sets are consistent. Thus all of the data will be utilized for this purpose. Lab 2 data are audits or checks on the Lab 1 test data and to be used to determine if the test data are valid based on the following three criteria and problem types:

(1) From past experience a maximum difference between audit and test results of 1 mg has been suggested. What can one infer concerning the test data?

(2) No standard is available, but a statistical analysis is to be conducted to compare the two sets of data with significance level of 0.95 or a risk level of 0.05.

(3) Assume that the audit data are unbiased and that it is desired to report the results of the test data as an estimated bias and expected range of variation using "3σ".

G.2.1 Criterion (1)

Based on criterion (1) above, all of the differences are less than 1 mg/filter and hence the test data would be considered to be valid. This analysis does not check the suggested standard. This will be done with respect to the second criterion.

Suppose further that these 10 audits represent a random sample of 10 test measurements selected from 50 which are checked for validity. What can one infer about the set of 50 measurements? The answer to this question requires some further background in statistical sampling than that given in these appendices, including Appendix I. However, with appropriate tables on sampling[1] one can, for example, infer that:

> "there is 50 percent confidence that the percent of good test measurements exceeds 90 percent,"

or

> "there is 95 percent confidence that the percent of good test measurements exceeds 75 percent,"

These are examples of the types of statements that can be made based on just this one data set. As additional data are obtained, one's confidence in a given percent of good test data should increase if the test data are actually satisfactory.

G.2.2 Criterion (2)

In this case no prior information is assumed about the expected deviation between a test measurement and an audit value. Thus the comparison is made on the basis of the behavior of the two sets of data or, really, the differences in the corresponding data pairs. For example, one wishes to determine if there is a significant bias in the measurements, and secondly what is a reasonable difference or standard to suggest for acceptance of test data?

G.2.2.1 Sign Test[2]- One simple test of a significant bias is to check the sign of the differences. If all ten differences are negative, then one has considerable doubt about the lack of a bias as it would be expected that on the average five would be positive and five negative if no bias were present. The chances that all ten are of one sign is like flipping an unbiased coin ten times and obtaining ten heads; since this is a very small chance, one usually infers that there is a bias. In the example there are nine negative differences among ten. The chances of 9 or 10 negative or positive differences, if there were no bias present, is given by the computation[2],

$$2 \left[10 \left(\frac{1}{2}\right)^{10} + 1 \left(\frac{1}{2}\right)^{10} \right] = \frac{22}{1024} = 0.0215.$$

Since this is very small, say less than 0.05, it is inferred that one set of data is biased with respect to the other set of data. Note that either the test data (lab 1) or the audit data (lab 2) or both may be biased. Unless some outside check of the results is available, e.g. against some reference standard, it is not possible to assume that one data set is not biased and the other set is biased.

G.2.2.2 Paired t-test - A somewhat more sophisticated check on the bias is provided by a paired t-test[2,3] which is described in almost any elementary text on statistical techniques. This is a very useful test for comparing paired data sets obtained by making two related measurements on the same sample or equivalent samples under the same conditions except, for example, a change in the operator and/or instrument, etc. After taking the differences, the test is conducted ignoring the original data and using only the differences. The mean $\bar{D}$ and standard deviation of the differences s_D are obtained as was done in Appendix F, that is, including the outlier,

$$\bar{D} = -0.291$$

$$s_D = 0.22.$$

Using these values, a value of t is calculated as follows,

$$t = \frac{\bar{D} - 0}{s_D/\sqrt{n}} = \frac{-0.291}{0.22/\sqrt{10}} = 4.18.$$

(See Table G.1 for a computational form for t.) This t value is then checked against the value in Table E.1 to determine if it is unusually small or large. Assuming 95 percent confidence (or 5 percent risk) it is obvious that this value is larger in absolute value than expected and hence one infers a bias exists as was previously suggested by the sign test. Note that eliminating the outlier would decrease s_D, increase the magnitude of $\bar{D}$ and hence increase t, and result in the same conclusion.

Next consider the implications for a suggested standard for differences between a test measurement and an audit value for this particular analysis. In answering this question, the standard deviation of the difference, s_D = 0.22 mg/filter is a measure of the variation of the differences about their own

Table G.1 TWO SAMPLES
SPLIT SAMPLES OR DUPLICATES
COMPARISON OF METHODS, LABS, OR REPEAT MEASUREMENTS

Sample No.	X_1	X_2	Difference (D)	Sample No.	X_1	X_2	$X_1-X_2=D$
1	____	____	____	11	____	____	____
2	____	____	____	12	____	____	____
3	____	____	____	13	____	____	____
4	____	____	____	14	____	____	____
5	____	____	____	15	____	____	____
6	____	____	____	16	____	____	____
7	____	____	____	17	____	____	____
8	____	____	____	18	____	____	____
9	____	____	____	19	____	____	____
10				20			
Total = ΣD			____				____

Calculations:

Mean Difference = $\bar{D} = \Sigma D/n =$ ____________

Standard Deviation of Differences:

$\Sigma D^2 =$ ________

$(\Sigma D)^2/n =$ ________

Difference = ________

Divide by n-1

$s_D^2 =$ ________

$s_D =$ ________

$s_D/\sqrt{n} =$ ________

Is the mean difference equal to 0?

Calculate

$$t = \frac{\bar{D}-0}{s_{\bar{D}}} = \frac{\sqrt{n}\bar{D}}{s_D} = \_\_\_\_\_\_ = \_\_\_\_\_\_$$

Compare to tabulated t value with n-1 degrees of freedom (DF), (Table E.1) for the selected level of significance or risk (e.g. for n-1 = 9 DF and 95 percent level of significance t = 2.262).

mean difference. Hence $3s_D$ = 0.66 mg/filter can serve as an expected variation which would be exceeded a relatively small percentage of the time just as one would use 3σ limits in developing control chart limits. However, there are two limitations to this approach which must be considered in developing reasonable limits, (1) only ten data pairs (differences) were available and this does not meet the usual recommendations for setting control limits, say $n = 20$ would be a preferred number of values, and (2) the bias is not considered. In practice some bias between labs, audit and test values is reasonable and an acceptable magnitude of bias must be determined. To determine an acceptable level a number of further data sets like those given in the example must be analyzed.

G.2.3 Criterion (3)

In this case it is assumed further that the audit data (lab 2) are not biased, and that it is desired to present the test data (lab 1) in terms of the bias and variation. In this case, the bias is estimated to be 0.29 mg/filter, that is, the test data are biased low on the average by 0.29 mg/filter. Hence, based on this data for a single test value, it is inferred that the true measurement in mg/filter is given by the following:

Test measurement - Bias ± $3s_D$

or $X + 0.29 \pm 3(0.22)$

or between $X - 0.37$ and $X + 0.95$ mg/filter.

G.3 SUMMARY

In summary, some of the possible uses of test and audit data are described in this appendix. If standard reference samples were available, they could be used to audit measurements made by analytical methods and the lab biases and expected deviations determined. Interlaboratory tests aid in estimating the within-lab and among-lab variation. The use of these tests is described in Appendix K.

G.4 REFERENCES

1. Smith, F. and Nelson, A. C., Guidelines for Development of Quality Assurance Programs and Procedures, Final Report to EPA on Contract No. EPA-Durham 68-02-0598, August 1973.

2. Dixon, W. J. and Massey, F. J., Introduction to Statistical Analysis, McGraw-Hill Book Co., Inc., New York, 1951.

3. Youden, W. J., Statistical Methods for Chemists, John Wiley and Sons, Inc., New York, 1951.

APPENDIX H

CONTROL CHARTS

H.1 DESCRIPTION AND THEORY

The control chart provides a tool for distinguishing the pattern of indeterminate (random) variation from the determinate (assignable cause) variation. This technique displays data from a process or method in a form which graphically compares the variability of all test results with the average or expected variability of small groups of data - in effect, a graphical analysis of variance, and a comparison of the "within groups" variability versus the "between group" variability (See Figure H.1 for the pattern of variation of data in a specific example).

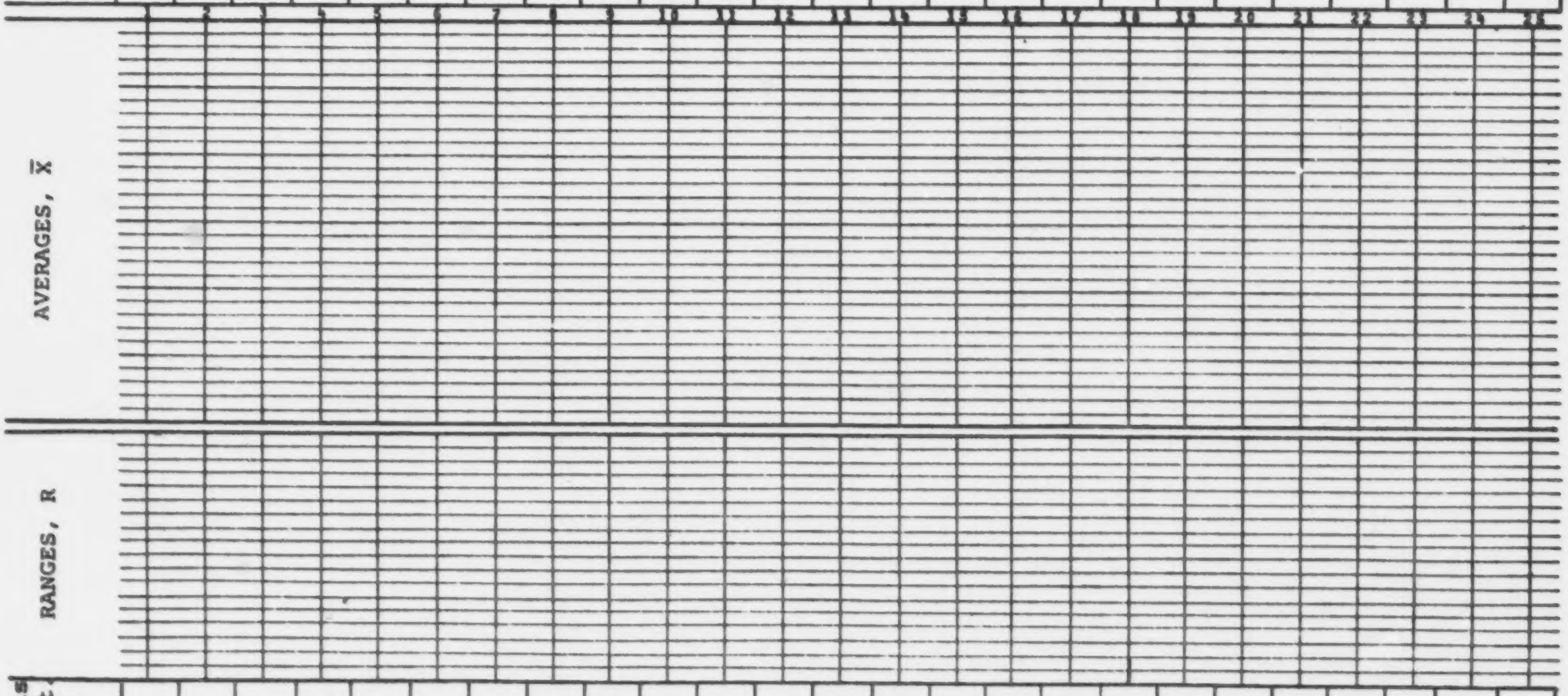
The data from a series of analytical trials can be plotted with the vertical scale in units of the test result and the horizontal scale in units of time or sequence of analyses. The average or mean value and the spread (dispersion or range) can be calculated. Figure H.1 is an example based on actual data obtained over 17 days as given in Table H.1. The data represent measurements of the same standard reference sample over the 17 days.

The control charts in this appendix are constructed on a standardized form. A blank copy of this form is included on the following page. The Handbook user should copy this form and use it for constructing control charts for all routine measurement systems. The important features of the standard form follow:

$\overline{X}$ and R CHART

PROJECT NAME ____________ MEASUREMENT PERFORMED ____________ MEASUREMENT UNITS ____________

DATE																											
MEASUREMENT	CODE	1																									
		2																									
		3																									
	RESULT	1																									
		2																									
		3																									
SUM																											
AVERAGE, $\overline{X}$																											
RANGE, R																											



Comments (Correct. Action, etc.)																									

Table H.1

MEASUREMENTS OF SO_2 CONCENTRATIONS[a] - BARIUM CHLORANILATE METHOD

Day	Duplicate Measurements, ppm		$\bar{X}$	R
1	29.2	22.7	25.95	6.5
2	28.4	25.2	26.80	3.2
3	29.2	26.4	27.80	2.8
4	32.9	----	32.90	---
5	27.9	30.2	29.05	2.3
6	26.4	31.8	29.10	5.4
7	31.8	31.5	31.65	0.3
8	39.4	29.1	34.25	10.3
9	28.6	29.2	28.90	0.6
10	28.0	26.2	27.10	1.8
11	31.2	35.2	33.20	4.0
12	37.6	31.8	34.70	5.8
13	26.9	29.0	27.95	2.1
14	30.7	28.0	29.35	2.7
15	31.9	26.8	29.35	5.1
16	28.9	36.2	32.55	7.3
17	27.8	31.4	29.60	3.6
Subtotals	516.8	470.7	-----	$\Sigma R = 63.8$

$\Sigma X = 987.5$

$\bar{X} = 29.92$ or 30 ppm $\qquad \bar{R} = 3.988$ or 4.0 ppm

$\theta = \bar{R}/d_2 = 3.988/1.13 = 3.53$

---

[a] Source: Parker, Carl D., Research Triangle Institute, NIOSH Report, Evaluation of Portable SO_2 Meters, April 1974.

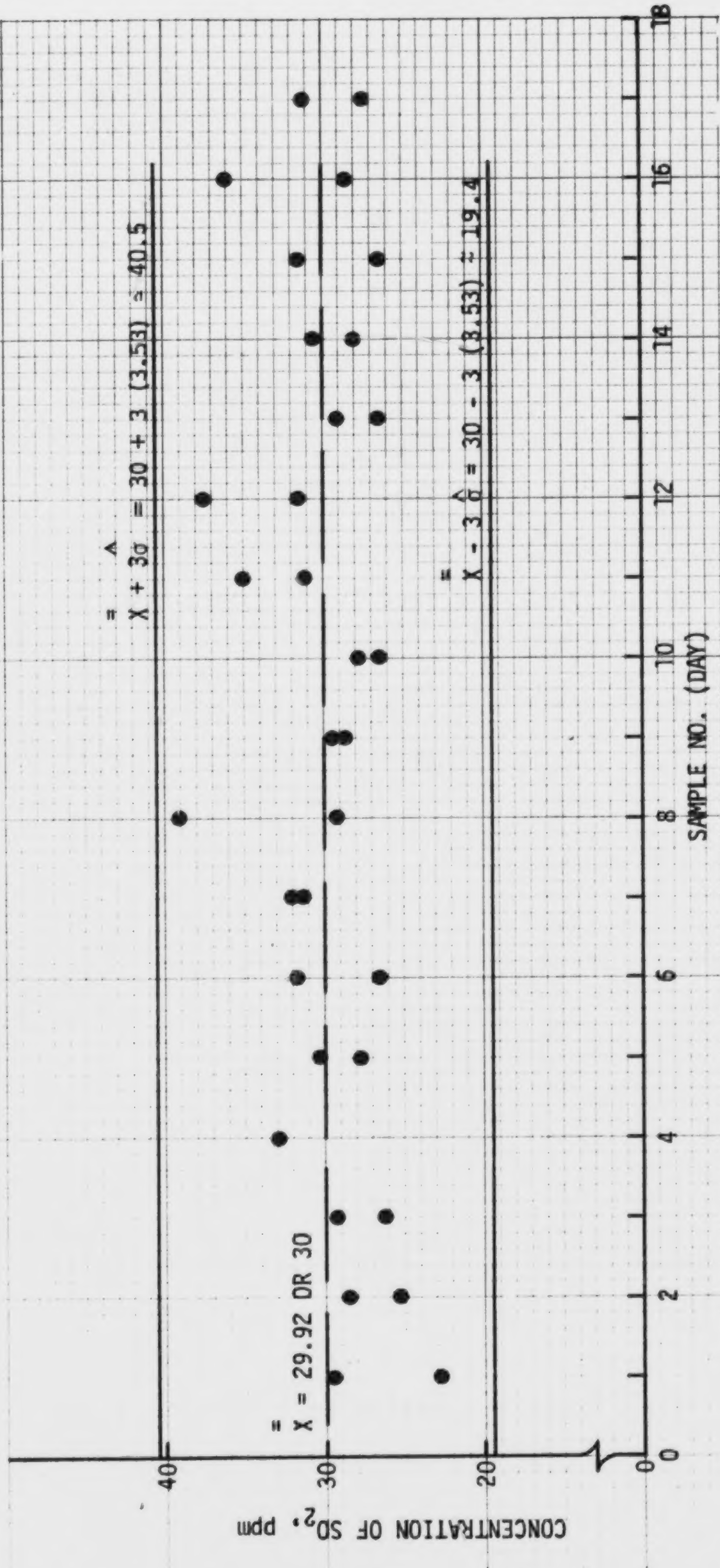


Figure H.1 Plot of data in Table H.1.

1. Measurement Performed - Record the pollutant, or parameter, measured and the method of measurement, e.g. SO_2 analysis of aqueous sodium sulfite standards,...Method.
2. Measurement Units - Metric Units.
3. Date - Write year next to the date. Write the month and day in the appropriate column.
4. Measurement Code - A number assigned to the measurement to permit easy reference to a more complete description of measurement conditions and results. This number, for example, should be traceable to an analyst notebook.
5. Measurement Results - Numerical results for the measurement code.
6. Comments - Note important observations and/or corrective actions taken, e.g. "instrument recalibrated". Comments should be entered when the measurement system is out of control and subsequent corrective action is taken.

The determination of appropriate control limits can be based on the capability of the procedure itself as known from past experience or on the specified requirements of the measurement procedure. Common practice sets the limits at the mean ± 3 standard deviations. Since the distribution of averages, and many distributions of individual values, exhibit a normal form, the probability of results falling outside the control limits can be readily calculated.

The control chart is actually a graphical presentation of quality control efficiency. If the procedure is "in control", the results will almost always fall within the established control limits. Further, the chart will disclose trends and cycles from assignable causes which can be corrected promptly. Chances of detecting small changes in the process average are improved when the average of several values is used for a single control point, (an $\overline{X}$ chart). As the sample size increases (for a single $\overline{X}$ point), the chance that small changes in the average will be detected is increased, provided the subgroup size is not altered as described in the following paragraph.

The basic procedure of the control chart is to compare "within group" variability to "between group" variability. For a single analyst running a procedure, the "within group" may well represent one day's output and the "between group" represents between days or day-to-day variability. When several analysts or several instruments or laboratories are involved, the selection of the subgroup unit is critical. Generally speaking, subgroups should be selected in a way that makes each subgroup as homogeneous as possible and that gives the maximum opportunity for variation from one subgroup to another. Assignable causes of variation should then show up as "between group" and not "within group" variability. Thus, if the differences between analysts may be assignable causes of variation, their results should not be lumped together in a "within group" subgrouping. The size of the subgroup is also important. Shewhart suggested 4 as the ideal size but subgroups of 2, 3, 5, or larger are often used.

H.2 APPLICATION AND LIMITATIONS

In order for quality control to provide a means for separating the determinate (systematic) from indeterminate (random) sources of variation, the analytical method must clearly emphasize those details which should be controlled to minimize variability. A check list would include:

Sampling procedures
Preservation of the sample
Aliquoting methods
Dilution techniques
Chemical or physical separations and purifications
Instrumental procedures
Calculating and reporting results

The next step to be considered is the application of control charts for evaluation and control of the more important of these unit operations. Decisions relative to the basis for construction of a chart are required:

(1) Select the variables (unit operations) to be measured.
(2) Choose method of measurement.
(3) Select the objective
    (a) Control of variability or precision
    (b) Control of accuracy of measurements
    (c) Control of completeness of reported data
    (d) Control of percentage of defective measurements.
(4) Select the size and frequency of subgroup samples:
    (a) Size--The analyst will often be dealing with samples of 2; process changes are detected more

quickly as the sample size is increased.

(b) Frequency of subgroup sampling--changes are detected more quickly as the sampling frequency is increased.

(5) Control limits can be calculated, but judgment must be exercised in determining whether or not the values obtained satisfy criteria established for the method, i.e. are the limits consistent with the solution or control of the problem? The control limits (CL's) can be calculated and control charts constructed as described in the following section.

H.3 CONSTRUCTION OF CONTROL CHARTS

H.3.1 Precision Control Charts

The use of range (R) in place of sample standard deviation (s) is justified for sample size $n \leq 10$ since R is nearly as efficient as s for use in estimating σ, and R is easier to calculate. The average range ($\bar{R}$) can be calculated from accumulated results, or from a known or assumed σ ($d_2\sigma$). Values of d_2 are tabulated vs. sample size n in Table H.2.

The steps employed in the construction of a precision control chart are given below and illustrated in Figure H.2, utilizing data on measurements of SO_2 concentrations in Table H.1.

(1) Calculate R for each set of analyses (subgroup)

(2) Calculate $\bar{R}$ from the sum of R values divided by the number (k) of sets (subgroups).

Table H.2

FACTORS FOR ESTIMATING THE STANDARD DEVIATION σ FROM THE RANGE R

Size of Sample	d_2	$\frac{1}{d_2}$
2	1.13	0.886
3	1.69	0.591
4	2.01	0.486
5	2.33	0.430
6	2.53	0.395
7	2.70	0.370
8	2.85	0.351
9	2.97	0.337
10	3.08	0.325
12	3.26	0.307
14	3.41	0.293

If $\hat{\sigma}$ is given, $\overline{R}$ can be calculated using $\overline{R} = d_2\hat{\sigma}$.

If $\overline{R}$ is known, an estimate of the standard deviation $\hat{\sigma} = \frac{1}{d_2} \times \overline{R}$.

Example H.1 If $n = 3$, $\hat{\sigma} = 5$,

$\overline{R} = 1.69(5) = 8.45.$

Example H.2 If $n = 2$, $\overline{R} = 3$

$\hat{\sigma} = 0.886(3) = 2.66.$

(3) Calculate the upper and lower control limits for the range:

$$UCL_R = D_4 \bar{R} , \quad LCL_R = D_3 \bar{R}$$

$$LCL_R = 0 \text{ when } n \leq 6$$

Since the analyses are in duplicates, $D_4 = 3.27$, $D_3=0$, see Table H.3, page 12.

(4) Calculate the upper warning limit:

$$UWL_R = \bar{R} + 2\sigma_R = \bar{R} + (2/3)(D_4 \bar{R} - \bar{R}) = D_5 \bar{R}$$

(from Table H.3, page 12).

(5) Chart $\bar{R}$, UWL_R and UCL_R on an appropriate scale which will permit addition of new results on a plot such as shown in Figure H.2.

(6) Plot results (R) and take action on out-of-control points. (See Section H.4).

H.3.2 Accuracy Control Charts - Mean or Nominal Value Basis

$\bar{X}$ charts simplify and render more exact the calculation of control limits since the distribution of data which conforms to the normal curve can be completely specified by $\bar{\bar{X}}$ ($\approx\mu$) and σ. Step by-step construction of an $\bar{X}$ (and R) control chart based on duplicate sets of results obtained from consecutive analyses of a control sample serves as an example (Bottom of Table H.3).

(1) Calculate $\bar{X}$ for each duplicate set and $\bar{\bar{X}}$ for all sets.

(2) Calculate R for each duplicate set and $\bar{R}$ for all sets.

(3) Calculate the upper and lower control limits by the equations:

$$UCL_{\bar{X}} = \bar{\bar{X}} + A_2\bar{R} , \quad LCL_{\bar{X}} = \bar{\bar{X}} - A_2\bar{R}$$

Values of A_2 vs. n are tabulated in Table H.3

Table H.3.

FACTORS FOR COMPUTING CONTROL CHART LINES[a]

Number of observations in subgroup, n	Factor A_2	Factor D_3	Factor D_4	Factor[b] D_5	Factor[c] D_6
2	1.88	0	3.27	2.51	0
3	1.02	0	2.57	2.05	0
4	0.73	0	2.28	1.85	0.15
5	0.58	0	2.11	1.74	0.26
6	0.48	0	2.00	1.67	0.33
7	0.42	0.08	1.92	1.62	0.38
8	0.37	0.14	1.86	1.58	0.42
9	0.34	0.18	1.82	1.55	0.45
10	0.31	0.22	1.78	1.52	0.48

[a] All factors in Table H.3 are based on the normal distribution.

[b] $D_5 = \frac{1 + 2D_4}{3}$ [c] $D_6 = \frac{5 - 2D_4}{3}$

Formulas for calculation	Example using data of Table H.1
$\bar{R} = \Sigma R \div k$	$\bar{R} = 63.8 \div 16 = 4.0$
$UCL_R = D_4\bar{R}$	$UCL_R = 3.27 \times 4.0 = 13.08$
$LCL_R = D_3\bar{R}$	$LCL_R = 0 \times 4.0 = 0$
$UWL_R = D_5\bar{R}$	$UWL_R = 2.51 \times 4 = 10.4$
$LWL_R = D_6\bar{R}$	$LWL_R = 0 \times 4.0 = 0$
$\bar{\bar{X}} = \Sigma\bar{X} \div k$	$\bar{\bar{X}} = \underline{\quad\quad} \div \underline{\quad\quad} = \underline{\quad\quad}$
or $\bar{\bar{X}}$[d] $= \Sigma X \div$ (total no. of measurements)	or $\bar{\bar{X}} = \underline{987.5} \div \underline{33} = \underline{29.92}$
$UCL_{\bar{X}} = \bar{\bar{X}} + A_2\bar{R}$	$UCL_{\bar{X}} = 29.92 + 1.88 \times 4 = 37.44$
$LCL_{\bar{X}} = \bar{\bar{X}} - A_2\bar{R}$	$LCL_{\bar{X}} = 29.92 - 1.88 \times 4 = 22.40$
$UWL_{\bar{X}} = \bar{\bar{X}} + \frac{2}{3} A_2\bar{R}$	$UWL_{\bar{X}} = 29.92 + 5.01 = 34.93$
$LWL_{\bar{X}} = \bar{\bar{X}} - \frac{2}{3} A_2\bar{R}$	$LWL_{\bar{X}} = 29.92 - 5.01 = 24.91$

[d] Use this second form when the numbers of measurements per group differ.

Table H.3.
(repeated with space for calculations)
FACTORS FOR COMPUTING CONTROL CHART LINES[a]

Number of observations in subgroup, n	Factor A_2	Factor D_3	Factor D_4	Factor[b] D_5	Factor[c] D_6
2	1.88	0	3.27	2.51	0
3	1.02	0	2.57	2.05	0
4	0.73	0	2.28	1.85	0.15
5	0.58	0	2.11	1.74	0.26
6	0.48	0	2.00	1.67	0.33
7	0.42	0.08	1.92	1.62	0.38
8	0.37	0.14	1.86	1.58	0.42
9	0.34	0.18	1.82	1.55	0.45
10	0.31	0.22	1.78	1.52	0.48

[a] All factors in Table H.3 are based on the normal distribution.

[b] $D_5 = \frac{1 + 2D_4}{3}$ [c] $D_6 = \frac{5 - 2D_4}{3}$

Formulas for calculations	Calculations:
$\bar{R} = \Sigma R \div k$ (number of subgroups)	$\bar{R} =$ ____ $\div$ ____ $=$ ____
$UCL_R = D_4\bar{R}$	$UCL_R =$ ____ x ____ $=$ ____
$LCL_R = D_3\bar{R}$	$LCL_R =$ ____ x ____ $=$ ____
$UWL_R = D_5\bar{R}$	$UWL_R =$ ____ x ____ $=$ ____
$LWL_R = D_6\bar{R}$	$LWL_R =$ ____ x ____ $=$ ____
$\bar{\bar{X}} = \Sigma\bar{X} \div k$	$\bar{\bar{X}} =$ ____ $\div$ ____ $=$ ____
or $\bar{\bar{X}}$[d] $= \Sigma X \div$ (total no. of measurements)	or $\bar{\bar{X}} =$ ____ $\div$ ____ $=$ ____
$UCL_{\bar{X}} = \bar{\bar{X}} + A_2\bar{R}$	$UCL_{\bar{X}} =$ ____ + ____ x ____ $=$ ____
$LCL_{\bar{X}} = \bar{\bar{X}} - A_2\bar{R}$	$LCL_{\bar{X}} =$ ____ - ____ x ____ $=$ ____
$UWL_{\bar{X}} = \bar{\bar{X}} + \frac{2}{3}A_2\bar{R}$	$UWL_{\bar{X}} =$ ____ + ____ $=$ ____
$LWL_{\bar{X}} = \bar{\bar{X}} - \frac{2}{3}A_2\bar{R}$	$LWL_{\bar{X}} =$ ____ - ____ $=$ ____

[d] Use this second form when the numbers of measurements per group differ.

(4) Calculate the upper and lower warning limits by the equations:

$$UWL_{\overline{X}} = \overline{\overline{X}} + (2/3)A_2\overline{R} \ , \quad LWL_{\overline{X}} = \overline{\overline{X}} - (2/3)A_2\overline{R}$$

(5) Construct the lines corresponding to $\overline{\overline{X}}$, $UCL_{\overline{X}}$, $LCL_{\overline{X}}$, and if desired, $UWL_{\overline{X}}$ and $LWL_{\overline{X}}$ as shown in Figure H.2.

(6) Plot $\overline{X}$ and take appropriate action on points which fall outside of the warning limits.

H.3.3 Control Charts for Percent of Defective Measurements

One type of control chart which is frequently used in industrial quality control is that for the fraction of defects in a production process. It is desired to maintain a low percentage of defects in order to minimize the expense of rework or waste. A standard or average percent of defects is established after examining the process for its past history simply by dividing the number of defects by the number of items produced, (or examined if less than 100 percent inspection).

In the case of a measurement process, one is concerned with the number of defective measurements among the total number of measurements to be taken. For example, suppose that n hi-volume filters are visually checked for defects and the number of defects is d, then the fraction of defects is $p = d/n$. Let the true or average fraction be p'. Then the limits can be obtained for p as follows:

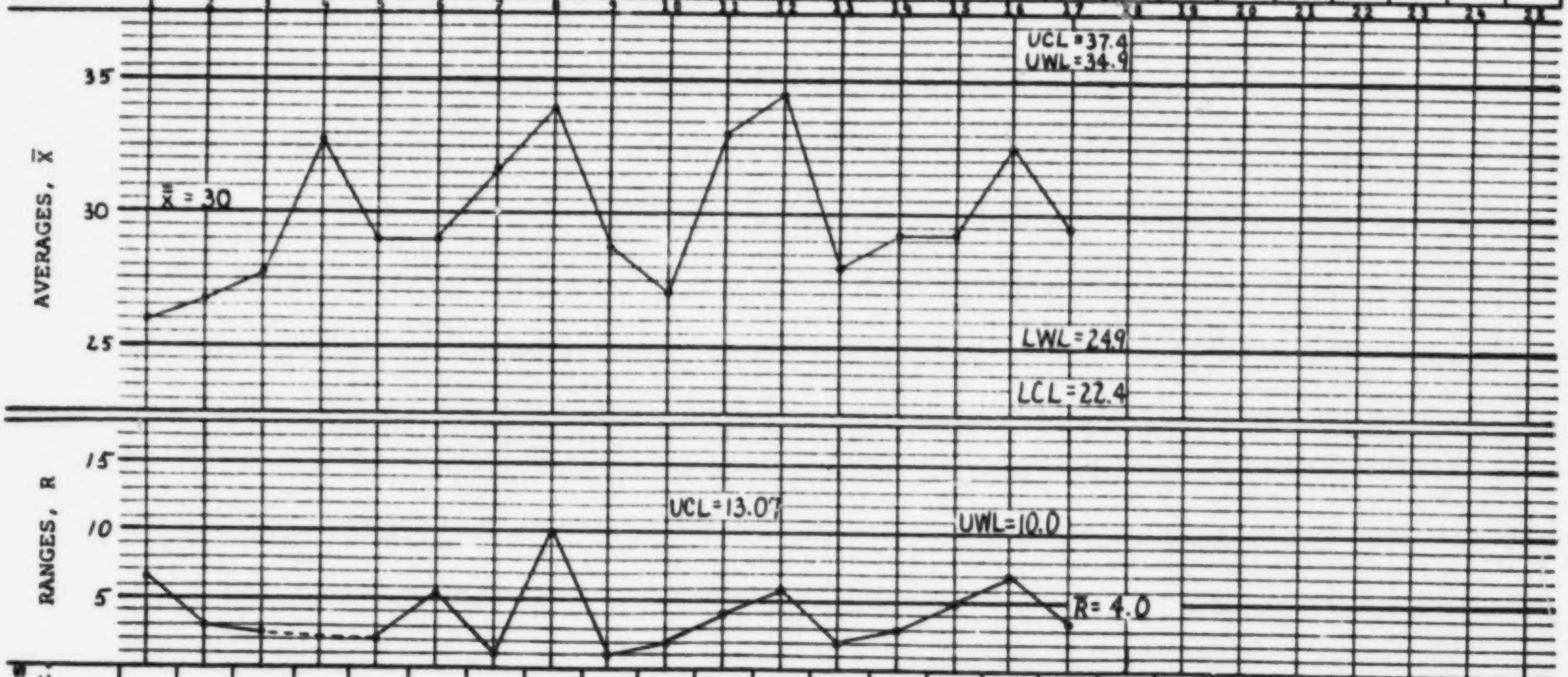
(1) Assume that p' has been determined on the basis of recent history of a process.

$\overline{X}$ and R CHART

PROJECT NAME Evaluation of portable SO_2 meters
MEASUREMENT SO_2 Concentration
PERFORMED Barium Chloranilate Method
MEASUREMENT UNITS PPM

1973

DATE			11/7	11/10	11/14	11/16	11/26	11/28	11/30	12/3	12/10	12/12	12/17	12/19	12/26	1/4	1/7	1/9	1/14								
MEASUREMENT	CODE	1	TASK ORDER NO.2																								
		2																									
		3																									
	RESULT	1	29.2	28.4	29.2	32.9	27.9	26.4	31.8	39.4	28.6	28.0	31.2	37.6	26.9	30.7	31.9	28.9	27.8								
		2	22.7	25.2	26.4	–	30.2	31.8	31.5	29.1	29.2	26.2	35.2	31.8	29.0	28.0	26.8	36.2	31.4								
		3																									
SUM			51.9	53.6	55.6	32.9	58.1	58.2	63.3	68.5	57.8	54.2	66.4	69.4	55.9	58.7	58.7	65.1	59.2								
AVERAGE, $\overline{X}$			26.0	26.8	27.8	32.9	29.0	29.1	31.6	34.2	28.9	27.1	33.2	34.7	28.0	29.4	29.4	32.6	29.6								
RANGE, R			6.5	3.2	2.8	0	2.3	5.4	0.3	10.3	0.6	1.8	4.0	5.8	2.1	2.7	5.1	7.3	3.6								



Comments (Correct. Action, etc.)

Figure H.2 Laboratory quality control worksheet.

(2) Calculate an upper control limit for p using

$$UCL_p = p' + 3 \left[\frac{p'(1-p')}{n}\right]^{1/2}.$$

(3) Calculate a lower control limit for p using

$$LCL_p = p' - 3 \left[\frac{p'(1-p')}{n}\right]^{1/2}.$$

(4) Draw lines $p = p'$, $p = UCL_p$, $p = LCL_p$ on a control chart with suitable scale for vertical axis to include expected range of variation of p'.

(5) Plot the fraction defective p for each sample along with the appropriate control limits. An example of such a control chart is given in Figure H.3 with the data taken from Table H.4. The sample size n is varied in order to indicate how such a control chart is constructed. A similar approach would be applicable to the $\bar{X}$ and R charts using appropriate n.[1]

This type of chart can also be used to control the completeness of reported data. The missing or incomplete data can be considered as the defects.

H.3.4 Control Charts for Individual Results

In many instances a rational basis for subgrouping may not be available, or the analysis may be so infrequent as to require action on the basis of individual results. In such cases, charts of individual values X are employed. However, the control limits should come from some subgrouping to obtain a measure of "within group" variability. This alternative has the advantage of displaying each result with respect to tolerance, or specification

Table H.4. COMPUTATION OF CONTROL LIMITS FOR FRACTION DEFECTIVE

$$p' = 0.05,\ \sigma_{100}\{p\} = \left[\frac{(0.05)(0.95)}{100}\right]^{1/2} \simeq 0.022,\ \sigma_{45}\{p\} = 0.0325,\ \sigma_{50}\{p\} = 0.0308$$

Sample no.	No. of measurements (n)	No. of defective measurements (d)	Fraction defective measurements (p=d/n)	UCL $p' + 3\sigma_n\{p\}$	LCL $p' - 3\sigma_n\{p\}$
1	50	1	0.02	0.142	0
2	50	3	0.06	0.142	0
3	45	2	0.044	0.148	0
4	100	6	0.06	0.116	0
5	45	4	0.088	0.148	0
6	50	3	0.06	0.142	0
7	45	0	0	0.148	0
8	45	3	0.066	0.148	0
9	50	2	0.04	0.142	0
10	100	7	0.07	0.116	0
11	45	2	0.044	0.148	0
12	45	1	0.022	0.148	0
13	45	3	0.066	0.148	0
14	50	4	0.08	0.142	0
15	45	1	0.022	0.148	0

$$\sigma_n\{p\} = \left[\frac{p'(1-p')}{n}\right]^{1/2}$$

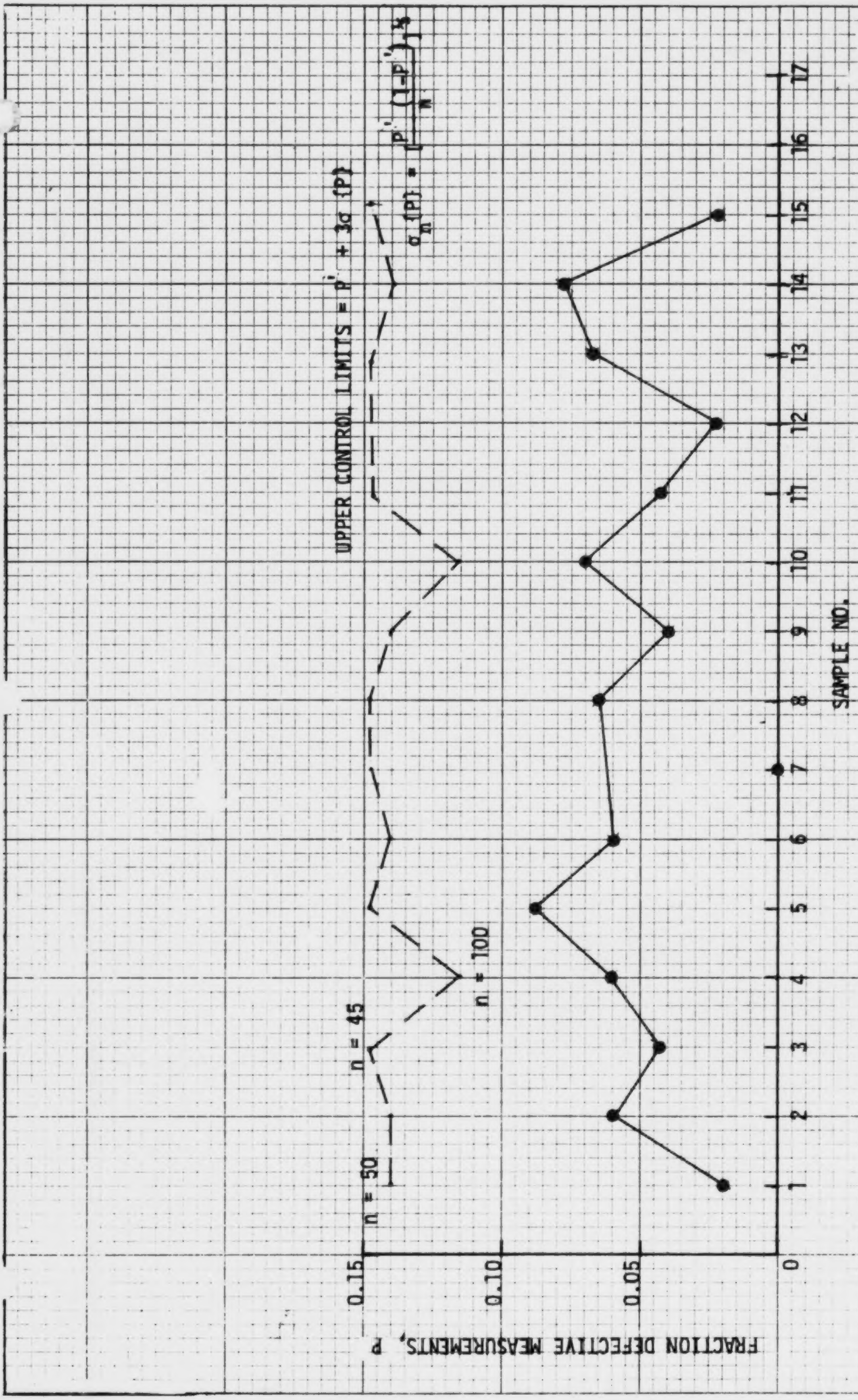


Figure H.3 Control chart for fraction of defective measurements.

limits (Figure H.1). The disadvantages must be recognized when considering this approach.

(1) The chart does not respond quickly to changes in the average.

(2) Changes in dispersion are not detected unless an R chart is included using the moving range, see Sec.H.3.5.

(3) The distribution of results must be approximately normal for the control limits to be valid.

H.3.5 Moving Averages and Ranges

The $\overline{X}$ control chart is more efficient than an X chart for individual values for moderate changes in the average as the subgroup size increases. A logical compromise between the X and $\overline{X}$ approach would be application of the moving average. The moving averages are obtained for the data given in the first column of Table H.1, and presented in Table H.5.

The moving average and range of two observations are obtained for each successive pair; that is, observations 1 and 2, 2 and 3, etc.

The moving range serves well as a measure of variation when no rational basis for subgrouping is available or when results are infrequent or expensive to gather.

Control charts can be constructed with the use of moving averages and ranges in the same manner as for ordinary charts. That is, one computes the grand average $\overline{\overline{X}}$ and $\overline{R}$ and then computes the limits.

Table H.5 MOVING AVERAGE AND RANGE TABLE (n=2)

Sample no.	Value	Moving averages of 2	Moving range
1	29.2	---	---
2	28.4	28.80	0.8
3	29.2	28.80	0.8
4	32.9	31.05	3.7
5	27.9	30.40	5.0
6	26.4	27.15	1.5
7	31.8	29.10	5.4
8	39.4	35.60	7.6
9	28.6	34.00	10.8
10	28.0	28.30	0.6
11	31.2	29.60	3.2
12	37.6	34.40	6.4
13	26.9	32.25	10.7
14	30.7	28.80	3.8
15	31.9	31.30	1.2
16	28.9	30.40	3.0
17	27.8	28.35	1.1
	Totals:	488.3	65.6
		$\bar{\bar{X}} = 30.52$	$\bar{R} = 4.1$

Refer to Table H.3

$UCL_R = D_4\bar{R} = 3.27 \times 4.1 = 13.4$

$LCL_R = D_3\bar{R} = 0 \times 4.1 = 0$

$UCL_{\bar{\bar{X}}} = \bar{\bar{X}} + A_2\bar{R} = 30.5 + 1.88 \times 4.1 = 38.2$

$LCL_{\bar{\bar{X}}} = \bar{\bar{X}} - A_2\bar{R} = 30.5 - 1.88 \times 4.1 = 22.8$

These limits are plotted on Figure H.4.
Warning limits could be computed in a similar manner.

$\bar{X}$ Chart (moving average of two consecutive values)

$$L\,C\,L_{\bar{X}} = \bar{\bar{X}} - A_2\bar{R}\ ,\ A_2 = 1.88 \text{ for } n = 2$$

$$U\,C\,L_{\bar{X}} = \bar{\bar{X}} + A_2\bar{R}$$

Range (R) Chart (moving range of two consecutive values)

$$L\,C\,L_R = D_3\bar{R} \qquad D_3 = 0 \text{ for } n = 2$$

$$U\,C\,L_R = D_4\bar{R} \qquad D_4 = 3.27 \text{ for } n = 2$$

The interpretation of the first point "out of control" is the same as for an ordinary chart. However, because successive plotted points use a common value, two consecutive points out of control on a moving average or range chart can be due to a single value out of control on an X chart where each measurement corresponds to only one point.

For further details on the moving average charts one is referred to Grant and Leavenworth[1], (pages 177 to 182). Because of their potential importance to air pollution measurements, an example application of the construction of such charts is given in Figure H.4 for moving averages and ranges of two consecutive observations.

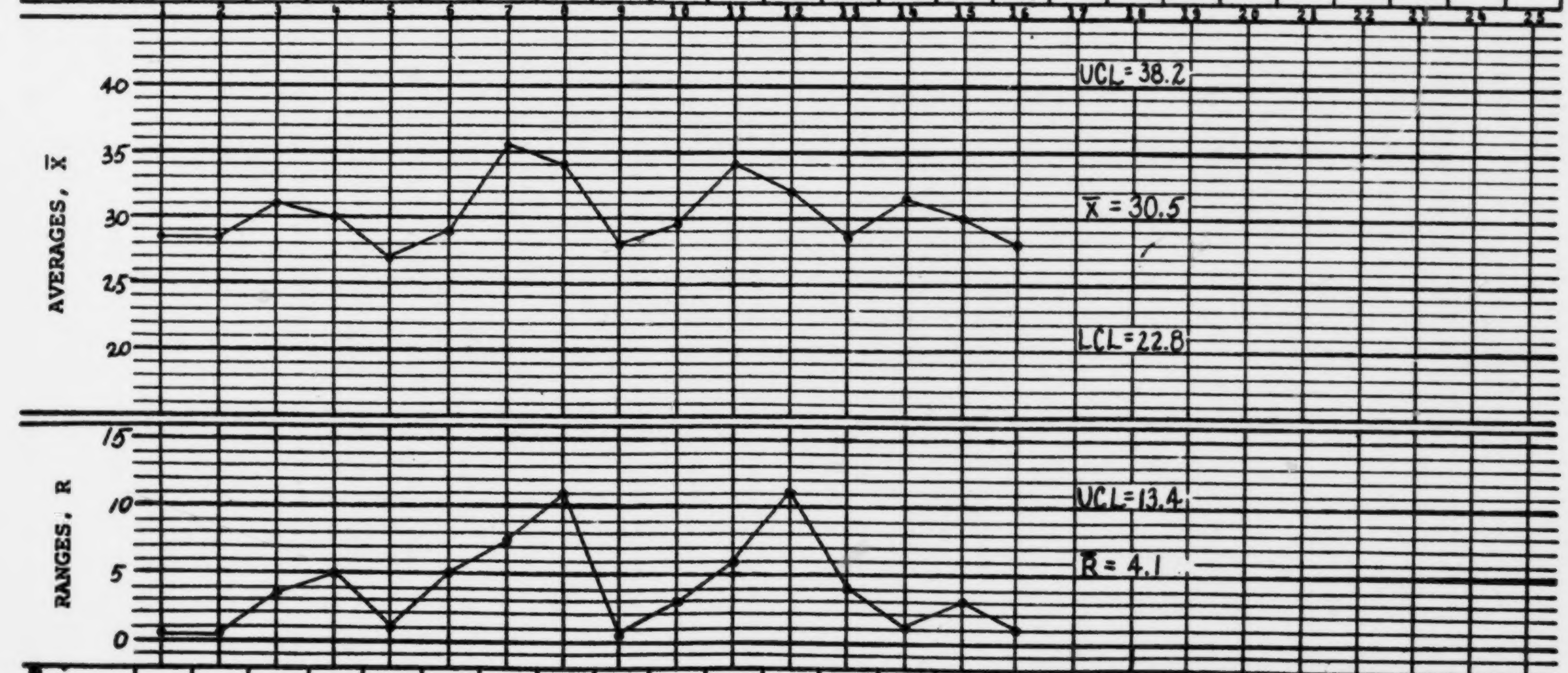
H.3.6 Other Control Charts for Variables

Although the standard $\bar{X}$ and R control charts for variables are the most common, they do not always do the best job. Several examples follow where other charts are more applicable.

$\overline{X}$ and R CHART

PROJECT NAME Evaluation of portable SO_2 meters
MEASUREMENT SO_2 Concentration
MEASUREMENT PERFORMED Barium Chloranilate Method
MEASUREMENT UNITS PPM

DATE		1973 11/7	11/10	11/14	11/16	11/26	11/28	11/30	12/3	12/10	12/12	12/17	12/19	12/26	1974 1/4	1/7	1/9
MEASUREMENT CODE	1	TASK ORDER NO. 2															
	2																
	3																
MEASUREMENT RESULT	1	29.2	28.4	29.2	32.9	27.9	26.4	31.8	39.4	28.6	28.0	31.2	37.6	26.9	30.7	31.9	28.9
	2	28.4	29.2	32.9	27.9	26.4	31.8	39.4	28.6	28.0	31.2	37.6	26.9	30.7	31.9	28.9	27.8
	3																
SUM		57.6	57.6	62.1	60.8	54.3	58.2	71.2	68.0	56.6	59.2	68.8	64.5	57.6	62.6	60.8	56.7
AVERAGE, $\overline{X}$		28.8	28.8	31.0	30.4	27.2	29.1	35.6	34.0	28.3	29.6	34.4	32.2	28.8	31.3	30.4	28.4
RANGE, R		0.8	0.8	3.7	5.0	1.5	5.4	7.6	10.8	0.6	3.2	6.4	10.7	3.8	1.2	3.0	1.1



Comments (Correct Action, etc.)

Figure H.4 Control chart for moving ranges for data in ... H.5.

H.3.6.1 Variable Subgroup Size - The standard $\bar{X}$ and R charts are applicable for a constant size subgroup on n. If n varies control limit values must be calculated for each sample size. Plotting is done in the usual manner with the control limits drawn in for each subgroup depending on its size. An example with varying control limits was described under control charts for fraction-defective measurements.

H.3.6.2 σ is Function of the Mean - When the dispersion is a function of concentration, control limits can be expressed in terms of a percentage of the mean. In practice such control limits would be given as in the example below.

± 5 units/liter for 0-100 units/liter concentration

± 5% for > 100 units/liter concentration

An alternative procedure involves transformation of the data. For example, logarithms would be the appropriate transformation when the standard deviation is proportional to the mean.

The frequent use of the coefficient of variation (CV) and %D* in air pollution measures suggests that it may be desirable to construct control charts for the sample $\hat{CV} = s/\bar{X}$ rather than transform the data. Using the results of Iglewicz and Myers[2] in comparing several approximations, the preferred method for obtaining percentiles of the distribution of $\hat{CV}_n$ is to use the result that

---

* %D is discussed in Appendix C and is referred to as percent error and is given by $\%D = \frac{100|X_1-X_2|}{(X_1+X_2)/2}$.
For n = 2, $\%D = \sqrt{2}\,\hat{CV}$ and thus the control limits for %D can be obtained by multiplying the control limits for $\hat{CV}$ by $\sqrt{2}$.

$$B \left[\frac{(n-1)\ \hat{CV}_n^2}{n+(n-1)\ \hat{CV}_n^2} \right] \approx \chi^2_{n-1}$$

where

$$B = n \left[1 + \frac{1}{CV^2} \right]$$

and

$$CV = \frac{\sigma}{\mu} = \text{population CV.}$$

See Reference 2 for definitions used in the above. To simplify the calculation of the limits, Figure H.5 contains the appropriate factors for samples sizes n = 2, 4 and 7.

As an example of how the limits may be calculated for a control chart for the coefficient of variation, $\hat{CV} = s/\overline{X}$, suppose that the CV is known to be 0.04 when the method is in a state of control. For a sample of size two (n = 2), the 99.5th percentile for the sample CV (= $s/\overline{X}$) is read from Figure H.5 to be about 0.113 or 11.3 percent. Hence, a control chart can be constructed with mean line at 0.04 or 4 percent and the upper control limit at 0.113 or 11.3 percent. If an observed CV exceeds 0.113 it is indicative of a possible lack of control of the relative standard deviation of the measurement.

If one is using the percent error measurement, i.e.

$$\%D = \frac{|X_1 + X_2| \times 100}{(X_1 + X_2)/2}$$

for n = 2, then $\%D = \sqrt{2}\ \hat{CV}$. Hence the limits are multiplied by $\sqrt{2}$ or the upper 99.5 percentile for %D for n = 2 and CV = 0.04,

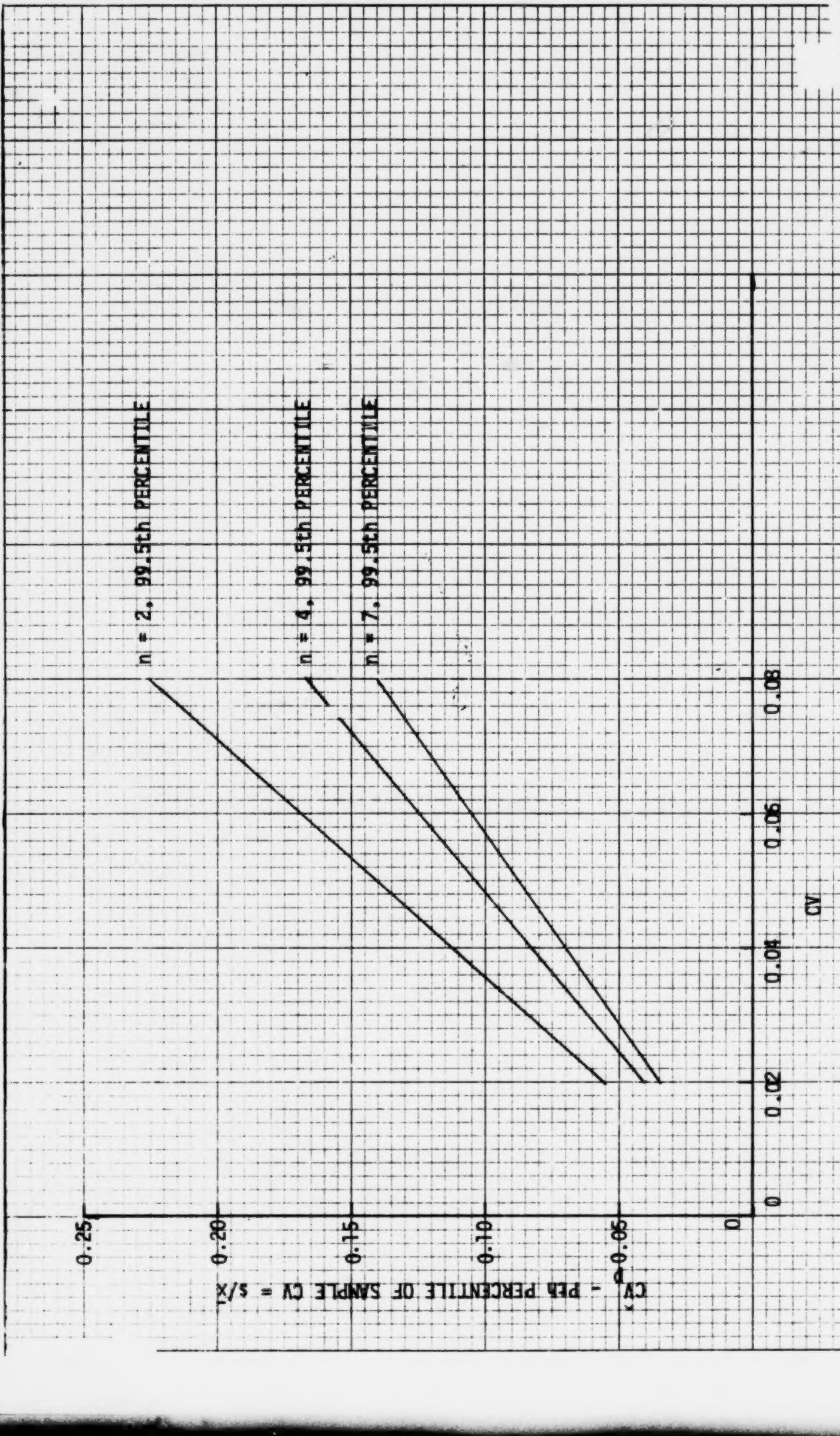


Figure H.5 Limits for Use in Constructing Control Charts for CV.

becomes 0.160 or 16 percent and the mean value would be $0.04 \times \sqrt{2} = 0.056$ or 5.6 percent. The 99.5 percentile is used as an approximation of the upper 3σ control limit which corresponds to the 99.86th percentile.

H.3.6.3. R or σ Charts - If the subgroup size exceeds 10, the range chart becomes inefficient. The use of a σ chart would then be appropriate. Where the cost of obtaining the test data is high, the increase in efficiency using σ rather than R may be worthwhile. Factors for constructing the chart for σ are given in Table H.6.

H.4 INTERPRETATION OF CONTROL CHARTS FOR OUT-OF-CONTROL

Various criteria have been used[3] to determine when the measurement system is out-of-control. The more important criteria for out-of-control are as follows:

(1) One or more points outside the control limits (3σ).

(2) A run of 2 or more points outside the warning limits (2σ).

(3) A run of 7 or more points. This might be a run up or run down or simply a run above or below the central line ($\bar{\bar{X}}$) on the control chart.

(4) Cycles or non-random patterns in the data. Such patterns may be of great help to the experienced operator.[1]

Table H.6*

FACTORS* FOR COMPUTING CONTROL CHART LINES

Based on $\bar{s}$

Number of observations in subgroup n	Factor for $\bar{X}$ Chart A_1	Factors for s chart	
		Lower control limit B_3	Upper control limit B_4
2	2.66	0	3.27
3	1.95	0	2.57
4	1.63	0	2.27
5	1.43	0	2.09
6	1.29	0.03	1.97
7	1.19	0.12	1.88
8	1.09	0.19	1.81
9	1.03	0.24	1.76
10	0.98	0.28	1.72
11	0.92	0.32	1.68
12	0.89	0.35	1.65
13	0.85	0.38	1.62
14	0.82	0.41	1.59
15	0.79	0.43	1.57
16	0.76	0.45	1.55
17	0.74	0.47	1.53
18	0.72	0.48	1.52
19	0.70	0.50	1.50
20	0.68	0.51	1.49

$$UCL_{\bar{X}} = \bar{\bar{X}} + A_1\bar{s}$$

$$LCL_{\bar{X}} = \bar{\bar{X}} - A_1\bar{s}$$

$$UCL_s = B_4\bar{s}$$

$$LCL_s = B_3\bar{s}$$

$\bar{s}$ = average standard deviation for k groups of n observations each, i.e. $\bar{s} = \Sigma s/k$

*All factors in Table H.6 are based on the normal distribution

Source: Grant, E. I. and Leavenworth, R. S., Statistical Quality Control, 4th Edition, McGraw-Hill Book Co., New York, p. 646, part of Table D.

$A_1 = \sqrt{\frac{n-1}{n}}\, A_1$ (as given in Table from Grant and Leavenworth)

H.5 STEPS IN DEVELOPING AND USING A CONTROL CHART SYSTEM

The following is the logic sequence which could be followed in developing and using a control chart system.

(1) Determine which key data to chart. Obviously, all available data cannot be plotted. Only the more important data should be plotted, such as results of calibrations, checks of standards or blinds, duplicate checks, etc.

(2) Decide what statistic to plot.

  (a) Control of average, plot $\bar{X}$ or X.

  (b) Control of variability, plot R or s.

(3) Evaluate the form of the distribution. Determine the form of the distribution, i.e., whether the distribution is normal, log-normal or as otherwise assumed.

(4) Eliminate any outliers from the sample of past data. Obviously, the existence of out-of-control points should not be used in the establishment of control limits.

(5) Determine the warning limits (2-sigma) and control limits (3-sigma).

(6) Use control chart form on page 2 and prepare any additional instructions for recording and plotting data and for taking action when out-of-control conditions are indicated.

(7) Draw in central line and control limits with bold lines. Where specification limits exist, these may also be shown.

(8) Maintain charts in the working area, if possible. Where possible and practicable, the operator should record the information and data, perform the necessary computations and plot the charts. (This would not be feasible when the results are for 'blinds' or unknowns to the operator.) In many cases the charts should be posted on a wall or otherwise kept in view in the working area. Some charts may be kept in a loose-leaf binder by the supervisor.

(9) Plot the points in bold fashion and join adjacent points by a straight line. It is very important that data be plotted as soon as it becomes available in order that out-of-control conditions can be detected as early as possible and that timely corrective action can be taken.

(10) Circle or otherwise highlight out-of-control points or conditions. It is also desirable to indicate visible trends on the charts.

(11) Take appropriate corrective actions when out-of-control conditions are indicated. Record on the chart the nature (and time) of the corrective action.

(12) Revise control limits periodically when indicated. When recent past history indicates an improvement in control, the revised limits will be tighter than previously. Good justification should be made before relaxing limits.

(13) Maintain an historical file:

(a) Data and computations used for determining control limits.

(b) Plotted control charts.

H.6 REFERENCES

1. Grant, E. I. and Leavenworth, R. S., Statistical Quality Control, Fourth Edition, McGraw-Hill Book Co., New York, 1972.

2. Iglewicz, B. and Meyers, R. H., Comparison of Approximations to the Percentage Points of the Sample Coefficients of Variation, Technometrics, 12 1:166-170. February 1970.

3. Duncan, A. J. Quality Control and Industrial Statistics, Third Edition, Richard D. Irwin, Inc., Homewood, Illinois, 1965.

4. Industrial Hygiene Laboratory Accreditations 589, NIOSH Quality Assurance Training Manual.

APPENDIX I

STATISTICAL SAMPLING

I.1 CONCEPTS

Suppose that filters to be used in high volume samplers are received in lots of size N = 100 and that it is desired to use filters with pH between two specified values, e.g.

$$6.5 \leq \text{pH} \leq 7.5.$$

If the pH test destroys a filter then it is necessary to employ some sampling procedure to determine whether the lot of N = 100 filters should be accepted for use consistent with prescribed specification limits such as given above. A sample of n filters is selected at random (this will be discussed in the next section), and if the number of defective filters, d, i.e. with pH below 6.5 or above 7.5, for example, exceeds a preselected value, c, the lot of filters is rejected. That is, it is presumed that the quality of the lot is not consistent with the desired specifications. This procedure is referred to in statistical and quality control texts as acceptance sampling by attributes. The approach can be applied to the acceptance of any product ordered in lots, and subject to desired specification levels. In some cases as above, 100 percent inspection is obviously imposible; while in other cases of non-destructive testing, 100 percent inspection is possible but may not be practicable because of costs or inspection fatigue and consequently an increased risk of the inspector mis-classifying the item.

I.2 HOW DOES ONE SELECT A RANDOM SAMPLE, STRATIFIED RANDOM SAMPLE?

I.2.1 Random Sample

Assume that a lot of N = 100 items is received, and a random sample of size n = 10 is to be drawn. First of all, what does one mean by a random sample of n = 10? This can be defined as a sample of 10 items selected in such a manner that every possible sample of size 10 has an equal chance of being selected. The sample can be selected with replacement (putting an item back after selection and inspection) or without replacement. Both cases will be considered. A strictly random procedure is as follows:

(1) Number the items in the lot from 1 to 100 (this does not have to be done by marking, but merely assign an ordering of the items to be sampled. e.g. if they are in a single stack the top item can be taken as 1, and the bottom item as 100 (Corresponding to 00 in the random number table).

(2) Use a table of random numbers or a random number generator such as a deck of cards numbered 0, 1, 2,...9.

(3) From the table of random numbers select a two digit number at random, often done by closing

Table I-1. SHORT TABLE OF RANDOM NUMBERS

05	57	23	06	26	23	08	66	16	11	75	28	81	56	14	62	82	45	65	80	36	02	76	56	[illegible]
37	78	16	06	57	12	46	22	90	97	78	67	39	06	63	60	51	02	07	16	75	12	90	41	16
23	71	15	08	82	64	87	29	01	20	46	72	05	80	19	27	47	15	76	51	58	67	06	80	54
42	67	98	41	67	44	28	71	45	08	19	47	76	30	26	72	33	69	92	51	95	23	26	85	76
05	83	03	84	32	62	83	27	48	83	09	19	84	90	20	20	50	87	74	93	51	62	10	23	30
60	46	18	41	23	74	73	51	72	90	40	52	95	41	20	89	48	98	27	38	81	33	83	82	94
32	80	64	75	91	98	09	40	64	89	29	99	46	35	69	91	50	73	75	92	90	56	82	93	24
79	86	53	77	78	06	62	37	48	82	71	00	78	21	65	65	88	45	82	44	78	93	22	78	09
45	13	23	32	01	09	46	36	43	66	37	15	35	04	88	79	83	53	19	13	91	59	81	81	87
20	60	97	48	21	41	84	22	72	77	99	81	83	30	46	15	90	26	51	73	66	34	99	40	60
67	91	44	83	43	25	56	33	28	80	99	53	27	56	19	80	76	32	53	95	07	53	09	61	98
86	50	76	93	86	35	68	45	37	83	47	44	92	57	66	59	64	16	48	39	26	94	54	66	40
66	73	38	38	23	36	10	95	16	01	10	01	59	71	55	99	24	88	31	41	00	73	13	80	62
55	11	50	29	17	73	97	04	20	39	20	22	71	11	43	00	15	10	12	35	09	11	00	89	05
23	54	33	87	92	92	04	49	73	96	57	53	57	08	93	09	69	87	83	07	46	39	50	37	85
41	48	67	79	44	57	40	29	10	34	58	63	51	18	07	41	02	39	79	14	40	68	10	01	61
03	97	71	72	43	27	36	24	$\boxed{59}$	88	82	87	26	31	11	44	28	58	99	47	83	21	35	22	88
90	24	83	48	07	41	56	68	11	14	77	75	48	68	08	90	89	63	87	00	06	18	63	21	91
98	98	97	42	27	11	80	51	13	13	03	42	91	14	51	22	15	48	67	52	09	40	34	60	85
74	20	94	21	49	96	51	69	99	85	43	76	55	81	36	11	88	68	32	43	08	14	78	05	34
94	67	48	87	11	84	00	85	93	56	43	99	21	74	84	13	56	41	90	96	30	04	19	68	78
58	18	84	82	71	23	66	33	19	25	65	17	90	84	24	91	75	36	14	83	86	22	70	86	89
31	47	28	24	88	49	28	69	78	62	23	45	53	38	78	65	87	44	91	93	91	62	76	09	20
45	62	31	06	70	92	73	27	83	57	15	64	40	57	56	(64)	42	35	40	93	55	82	08	78	87
31	49	87	12	27	41	07	91	72	64	63	42	06	66	82	71	28	36	45	31	99	01	03	35	76
69	37	22	23	46	10	75	83	62	94	44	65	46	23	65	71	69	20	89	12	16	56	61	70	41
93	67	21	56	98	42	52	53	14	86	24	70	25	18	23	23	56	24	03	86	11	06	46	10	23
77	56	18	37	01	32	20	18	70	79	20	85	77	89	28	17	77	15	52	47	15	30	35	12	75
37	07	47	79	60	75	24	15	31	63	25	93	27	66	19	53	52	49	98	45	12	12	06	00	32
72	08	71	01	73	46	39	60	37	58	22	25	20	84	30	02	03	62	68	58	38	04	06	89	94
55	22	48	46	72	50	14	24	47	67	84	37	32	84	82	64	97	13	69	86	20	09	80	46	75
69	24	98	90	70	29	34	25	33	23	12	69	90	50	38	93	84	32	28	96	03	65	70	90	12
01	86	77	18	21	91	66	11	84	65	48	75	26	94	51	40	51	53	36	39	77	69	06	25	07
51	40	94	06	80	61	34	28	46	28	11	48	48	94	60	65	06	63	71	06	19	35	05	32	56
58	78	02	85	80	29	67	27	44	07	67	23	20	28	22	62	97	59	62	13	41	72	70	71	07
33	75	88	51	00	33	56	15	84	34	28	50	16	65	12	81	56	43	54	14	63	37	74	97	59
58	60	37	45	62	09	95	93	16	59	35	22	91	78	04	97	98	80	20	04	38	93	13	92	30
72	13	12	95	32	87	99	32	83	65	40	17	92	57	22	68	98	79	16	23	53	56	56	07	47
22	21	13	16	10	52	57	71	40	49	95	25	55	36	95	57	25	25	77	05	38	05	62	57	77
97	94	83	67	90	68	74	88	17	22	38	01	04	33	49	38	47	57	61	87	15	39	43	87	00
09	03	68	53	63	29	27	31	66	53	39	34	88	87	04	35	80	69	52	74	99	16	52	01	65
29	95	61	42	65	05	72	27	28	18	09	85	24	59	46	03	91	55	38	62	51	71	47	37	38
81	96	78	90	47	41	38	36	33	25	05	90	26	72	85	23	23	30	70	51	56	93	23	84	80
44	62	20	81	21	57	57	85	00	47	26	10	87	22	45	72	03	51	75	23	38	38	56	77	97
68	91	12	15	08	02	18	74	56	79	21	53	63	41	77	15	07	39	87	11	19	25	62	19	30
29	33	77	60	29	09	25	09	42	28	07	15	40	67	56	29	58	75	84	06	19	54	31	16	53
54	13	39	19	29	64	97	73	71	61	78	03	24	02	93	86	69	76	74	28	08	98	84	08	23
75	16	85	64	64	93	85	68	08	84	15	41	57	84	45	11	70	13	17	60	47	80	10	13	00
36	47	17	08	79	03	92	85	18	42	95	48	27	37	99	98	81	94	44	72	06	95	42	31	17
29	61	08	21	91	28	76	72	84	98	26	23	66	54	86	88	95	14	82	57	17	99	16	28	99

Reproduced with permission from "A Million Random Diget", Rand Corp., Copywrite 1955, the Free Press.

one's eyes and placing the finger on the page and identifying the closest number pair just above the finger. See Table I.1 as an example.

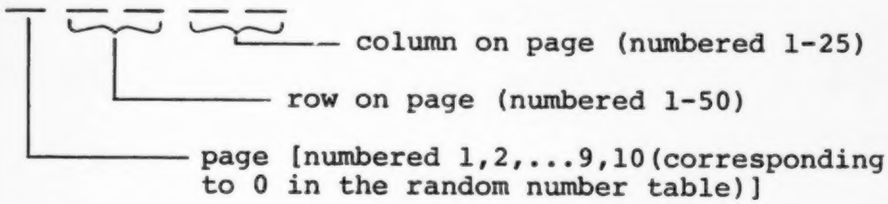
(4) Record the number obtained in (3) above and the following nine two-digit numbers, e.g. 59, 11, 13, 99, 93, 19, 78, 83, 72, 62. (See Table I.1 for 59 in a rectangular box).

(5) If there are any repeats, select another number in case of sampling without replacement; otherwise the ten numbers identify the items to be selected from the lot and checked against the specifications.

(6) If one wishes to use a deck of ten cards numbered 0 to 9 instead of a random number table, one card can be selected at random, after shuffle say, replaced and followed by a second drawing after a reshuffle. The two digits yield a two-digit number between 00 and 99 and hence determine the item to be selected from the lot.

(7) Repeat procedure in (6) to obtain nine additional two-digit numbers for the total of the items to be selected. If there are any repeats draw another number, etc.

There are more elaborate procedures for finding a starting point which are applicable when there are several pages in the table. For example, the pages and the rows and columns on a page may be numbered, then a random number is selected to identify which page, row, and column is used as a starting point. In Table I.1 there are 25 two digit columns, 50 rows, and suppose there were ten such pages similar to the page given. A five digit number could be drawn to select the page, row and column as suggested below.


column on page (numbered 1-25)
row on page (numbered 1-50)
page [numbered 1,2,...9,10(corresponding to 0 in the random number table)]

After using a set of random numbers, the stopping place can be noted, and the next set drawn with the next number in sequence. If the bottom of the table is reached, assuming the numbers are taken in vertical sequence, the next number can be taken from the top of the following column of numbers of the same number of digits. Theoretically, one can read the numbers horizontally if one wishes.

The example previously described involved the selection of a sample from 100 items and thus the numbering of the items can be put in one-to-one correspondence with the set of two digit numbers as follows:

01	02	...	...	99	00
1	2	...	...	99	100

If the lot or population from which the sample is drawn does not consist of 10, 100, 1000, etc. items, then the correspondence can often be altered to simplify the drawing of the sample rather than throwing out all of the numbers above N the size of the population. Two examples are given below, one in which N divides evenly into 100, and one for which this is not the case.

EXAMPLE I.1 Draw a random sample of seven days from 25 days using the random number Table I.1.

A random position in the table is first selected, for example, the number 54 in the circle. Starting at the point the following seven numbers are recorded:

No.	Random No.	Remainder
1	54	4
2	71	21
3	71	21
4	23	23
5	17	17
6	53	3
7	2	2
8	64	14

These numbers exceed 25 in five cases. Thus one can either (1) continue to select numbers until all numbers fall between 1 and 25 (or 00 to 24) or (2) divide each number which is larger than or equal to 25 by 25 and use the remainder as the random number. In this example one additional number had to be drawn to avoid repeats. The final sample consists of days numbered 2, 3, 4, 14, 17, 21, and 23.

In the example, the correspondence is established through the following:

Items in population:	1,	2,	3,	... ,	25
Nos. in table	00,	01,	02,	... ,	24
	25,	26,	27,	... ,	49
	50,	51,	52,	... ,	74
	75,	76,	77,	... ,	99

Example I.2 Draw a sample of n = five days from N = 15 days using the random number Table I.1.

In the previous example, 100 is exactly divisible by 25 and thus each number 0 to 24 has an equal chance of being drawn under the system given, i.e. dividing by 25 and taking the remainder. If this same system is used for N = 15, n = 5, numbers 00 through 09 would have a greater chance of being selected than the numbers 10 to 14 because of the number 90 to 99 being in the table. Thus disregarding these numbers the same approach can be employed as above. Using the same set of numbers for illustration the sample would be obtained as follows:

No.	Random No. (RN)	RN ÷ by 15 (Remainder)
1	54	9
2	71	11
3	71	11 (Repeat)
4	23	8
5	17	2
6	53	8 (Repeat)
7	02	2 (Repeat)
8	64	4

Thus the days numbered 2, 4, 8, 9, and 11 would be selected.

In this example, the correspondence is established as follows:

Items No. in Population:	1,	2,	3,	... ,	15
No. in Random Number Table:	00,	01,	02,	... ,	14
	15,	16,	17,	... ,	29
	30,	31,	32,	... ,	44
	45,	46,	47,	... ,	59
	60,	61,	62,	... ,	74
	75,	76,	77,	... ,	89
discarded	90,	91,	92,	... ,	99

I.2.2 Stratified Random Sample

Suppose the items are in stacks of 10 each. One could select one item at random from each stack of 10. The one item could be selected using either a random number table (one digit- 0-9) or the deck of cards numbered 0, 1, 2,...9. This is a stratified random sample. That is, the lot of items to be sampled is first stratified or subdivided into sublots and a random sample is selected from each sublot in accordance to its size relative to the entire lot.

In many applications in sampling, the strata are defined to coincide with some characteristics of the population to be sampled. For example, the strata may be items produced or manufactured on one day, shift, or from one lot of material. In the case of sampling days from a year the strata might be weeks, months, or seasons or coincide with known production schedules of a particular manufacturing plant. In a stratified random sample the strata are sampled proportionally, thus providing in many applications a more representative sampling of the population. This is particularly true when there is considerable variation among strata and little variation within strata.

I.2.3 Systematic Sample - (A Pseudo-Random Sample)

Another procedure is to select a systematic sample of items by selecting the first item at random and every tenth item thereafter, depending on the sample size. In the case of selecting n = 10 from N = 100 items, the first item is selected at random from the first 10 items, and suppose it is a three (3), then items 13, 23, 33, etc.,...93 are selected. One needs to assess the possible consequences of the alternate sampling schemes, for if there is any possible relationship between the defective items and the systematic selection, it is obvious that some biased results could occur. For example, if the sample collection of a pollutant coincided with meteorological cycles, the concentrations would tend to be less variable and may be relatively high (or low).

The reader is referred to standard statistical texts for a more complete discussion of sampling procedures.[1,2,4,5]

I.3 ACCEPTANCE SAMPLING BY ATTRIBUTES

Now consider the problem of determining whether a lot of items should be accepted on the basis of given specifications. Suppose that a defective item is one with a physical defect which can be identified by a visual test or one for which a physical measurement falls outside a prescribed value(s) as in the example of the pH of the filters. Thus the sampling is by attributes, i.e. an item is identified as either a defect or a good item (non-defect), often referred to as go/no-go inspection where this refers to whether the item meets the specification when checked by a gauge, calipers, sieve, etc.

A tabulation of sampling plans is given in MIL-STD-105.[1] A complete discussion of the plans is given therein and not repeated here.

I.4 ACCEPTANCE SAMPLING BY VARIABLES

A considerable savings in sampling may be achieved if the decisions concerning the acceptance of a lot of data (measurements) can be made on the basis of the actual measurement (a continuous value) rather than whether the measurements are outside specific limits. For example, in checking or auditing the quality of certain measurements the decision can be made to classify a routinely obtained field measurement as a defect if it deviates from the audit value by more than say 10 percent of the audited value. This results in sampling by attributes because one is classifying the items only as defects or nondefects (good measurements) rather than using the measured difference in the two values. If the latter is used, the decision on acceptance of a lot is made by variables, i.e. on the basis of the mean and standard deviation of the sample of n measurements and given constants, very much like that used in control charts.

Variable sampling plans are described and tabulated in MIL-STD-414 (normal distribution only).[4] The reader is referred to this standard for further details and some specific variable sampling plans.[3]

I.5 REFERENCES

1. MIL-STD-105, Sampling Procedures and Tables for Inspection by Attributes, Government Printing Office, Washington, D. C.

2. Juran, J. M., Quality Control Handbook, Third Edition, McGraw Hill Book Co., New York, 1974.

3. Owen, D. B., Variables Sampling Plans Based on the Normal Distribution, Technometrics, Vol. 9, No. 3, August 1967, pp. 417-424.

4. MIL-STD-414, Sampling Procedures and Tables for Inspection by Variables for Percent Defective, U.S. Department of Defense, Military Standard, Government Printing Office, Washington, D. C.

5. Bowker, A. H. and H. P. Goode, Sampling Inspection by Variables, McGraw Hill Book Co., New York, 1952.

APPENDIX J

CALIBRATION

J.1 CONCEPTS

One of the most important steps in the measurement process is the calibration of the instruments for use in measuring the concentration of pollutants in ambient air. In addition to a multipoint calibration performed periodically, there need to be frequent checks, such as zero-span checks, to determine if there is a significant change in the calibration curve obtained at the most recent multipoint calibration. These checks might be performed daily, or less frequently after considerable experience with the measurement process and with appropriate use of control charts. The practice of making daily adjustments of the equipment as a result of zero-span checks could lead to introducing more variability in the system rather than correcting the problems. It is best to make adjustments only after significant evidence has indicated that a recalibration should be made. The techniques for making some of the decisions concerning calibrations are not discussed in the general literature and some specific guidelines are presented in this section for use by the analyst/operator performing the calibrations. These guidelines emphasize that care should be taken to note trends in the results over time due to degradation of an instrument, calibration gas or other standards used in the frequent checks based on actual data. However, the operator should always consider his subjective feeling concerning a change in either the instrument/

standard and/or environment which may alter the calibration. He should make the checks that he considers necessary to assure himself that valid results will be obtained with the measurement system.

The discussion of calibration procedures requires some consideration of statistical techniques for fitting a linear or nonlinear function to a set of calibration data. The technique most frequently used is the least-squares method. Although calibration data can usually be adequately fitted by "eyeball", that is, fitting a line or curve to the data based on a subjective fit by an analyst using a straight edge or a french curve, this technique must be supplemented by some calculation of how well the data are "fitted" by the curve in order to predict the precision/accuracy of the reported data and to make objective checks on whether the calibration may have changed. The least-squares method and some associated statistical computations will be discussed briefly in Section J.2.

Some specific problems to be considered in this section are indicated below under the two headings, multipoint calibration and zero-span calibration.

Multipoint Calibration (MPC)

Some questions to consider with respect to MPC are:

(1) How often to repeat the calibration?

(2) How many points (levels of standards, e.g. concentrations of calibration gases) to use?

(3) How to space the levels of the standards?

(4) Is the calibration curve linear or nonlinear?

(5) Is the calibration equally precise for all concentration levels?

(6) How to estimate the precision of the estimates read from the calibration curve or table?

(7) How does the expected range of measured concentrations affect the selection of the levels to be used in the calibration?

Zero-Span Calibration (OSC)

Some of the questions to be considered for the OSC are:

(1) How often do we make an OSC?

(2) At which levels of the standard should the checks be made?

(3) What is the day-to-day variation among the checks?

(4) How is a quality control chart used in determining if a significant drift or change in the calibration curve has ocurred, based on the OSC?

(5) How does the expected range of concentrations affect the selection of the OSC levels?

There are additional questions which may be asked relative to the calibration procedure, but the above represent some of the most important ones. The answers to all of the above questions are not easy to discuss briefly in this section. The approach will be to use examples to indicate the approach and refer to some appropriate statistical texts (References 1, 2) for a more detailed discussion of the basic mathematical structure of the problem.

J.2 MULTIPOINT CALIBRATION (MPC)

Example J.1 provides multipoint calibration for an NO_2 analyzer for five levels of concentration. These data are plotted in Figure J.1 to give y as a function of X. A straight line was fitted to the five data points using the least-squares method. The necessary calculations are indicated below the example.

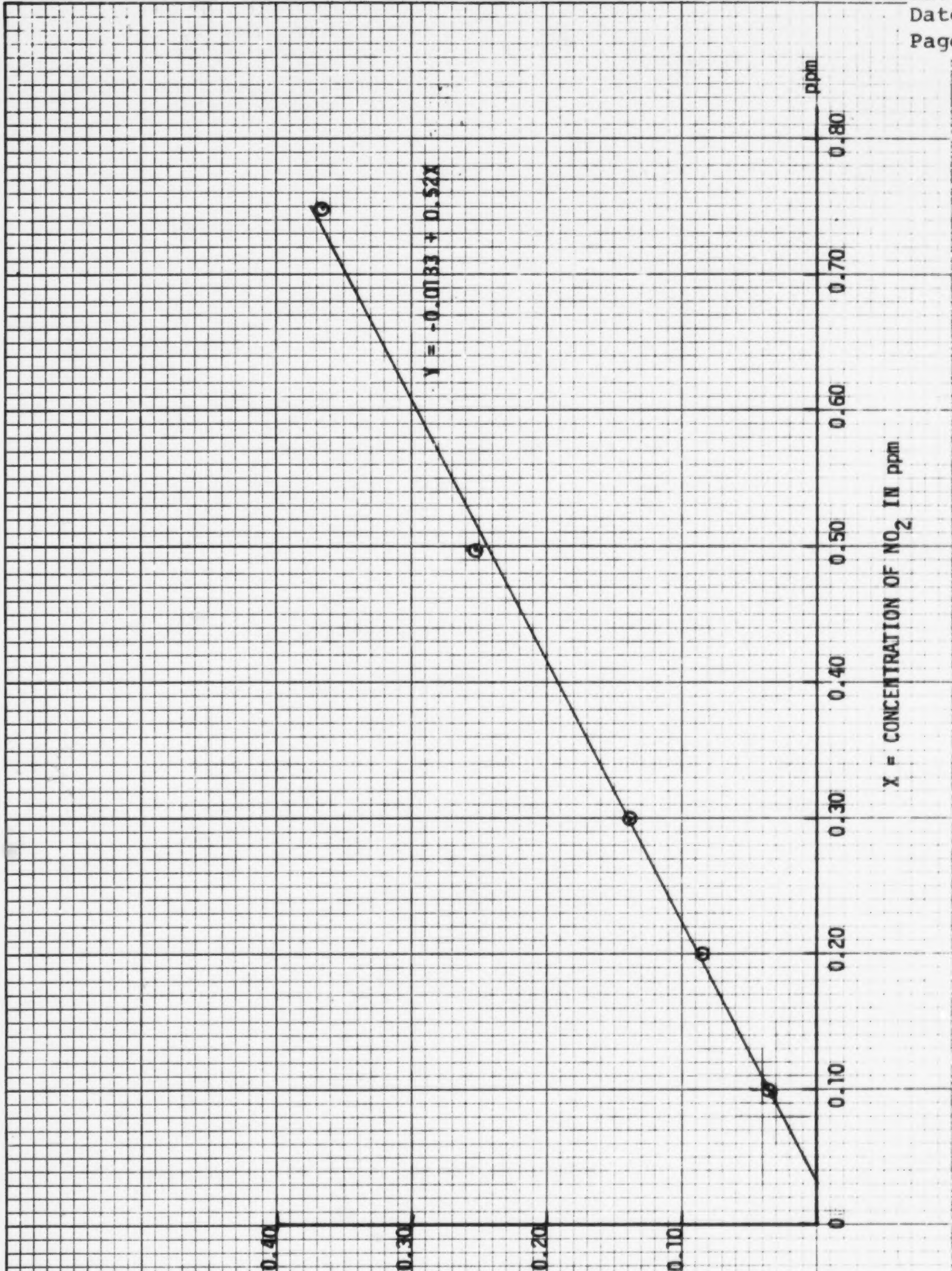


Figure J.1 Multipoint calibration of NO_2 analyzer.

Example J.1

X (Concentration of NO_2, ppm)	y (Analyzer Reading , volts)
0.10	0.039
0.20	0.086
0.30	0.140
0.50	0.254
0.75	0.369
Totals 1.85	0.888

Assuming that the calculations are performed manually using a desk (or portable) calculator without a completely programmed feature, the following steps are performed in obtaining the equation of the "best fit" line by the least-squares method and the variance of the responses about the fitted line.

Preliminary Computations

$$\Sigma X = 1.85 \qquad \Sigma y = 0.888$$

$$\Sigma X^2 = 0.9525 \qquad \Sigma y^2 = 0.2292$$

$$\Sigma Xy = 0.4668$$

The notation ΣX is an abbreviation for $\sum_{i=1}^{n} X_i$, similarly for Σy, ΣX^2, Σy^2 and ΣXy. The subscripts are dropped for ease in typing and reading.

$\overline{X} = \Sigma X/n = 1.85/5 = 0.37 \qquad \overline{y} = \Sigma y/n = 0.888/5 = 0.1776$

$$b = \frac{\Sigma Xy - (\Sigma X)(\Sigma y)/n}{\Sigma X^2 - (\Sigma X)^2/n}$$

$$b = \frac{0.4668 - (1.85)(0.888)/5}{0.9525 - (1.85)^2/5} = \frac{0.1383}{0.2680} = 0.516.$$

Record the numerator and denominator of b for future use.

The equation of the line fitted to the data is written as

$$\begin{aligned} Y &= \overline{y} + b(X-\overline{X}) = a + bX \\ &= 0.1776 + 0.516\ (X-0.37) \\ &= -0.013 + 0.52X \end{aligned}$$

where Y is the predicted mean response for the corresponding X.

The next step in the calculation is to determine how well the data are fitted by the line. The measure used is the sum of squares of deviations between the data points and the corresponding predicted Y given by the equation above, divided by the sample size less 2 (i.e. n-2). The reason for dividing by n-2 instead of n-1, as used in Appendix C, is that two parameters of the line have been estimated from the data (the intercept and the slope), thus the remaining degrees of freedom in fitting the five data points is n-2 = 3.

The calculation of the variance and standard deviation of the responses is performed as follows:

$$s^2_{y|X} = \frac{\Sigma y^2 - (\Sigma y)^2/n - b\{\Sigma Xy - (\Sigma X)(\Sigma y)/n\}}{n-2}$$

$$= \frac{0.2292 - (0.1776)(0.888) - (0.516)(0.1383)}{3}$$

$$s^2_{y|X} = 0.000127/3 = 0.000042$$

and $$s_{y|X} = 0.0065 \text{ volts.}$$

The notation $S_{y|X}$ is not used when it is clearly understood from the context of the discussion that the standard deviation is that of the response or y value about the regression curve (or line).

In the formula for $s^2_{y|X}$, all of the information has been previously calculated, Σy^2, $\bar{y}$, Σy, b, and the quantity in brackets {} is the numerator of the expression for b. The last term in the computation of the numerator of $s^2_{y|X}$ can also be written as

$$b \times b\{\Sigma X^2 - (\Sigma X)^2/n\} \text{ or } b^2\{\Sigma X^2 - (\Sigma X)^2/n\};$$

that is, b^2 multiplied by the denominator used in calculating b.

Line Through the Origin - If the line must pass through the origin (X = 0, Y = 0) then Y = b'X where

$$b' = \frac{\Sigma Xy}{\Sigma X^2} = \frac{0.4668}{0.9525} = 0.49,$$

compared to 0.516 when not forced through the origin.

The estimated variance $s^2_{y|X}$ for the case in which the line must pass through the origin is given by

$$s^2_{y|X} = \frac{\sum_{i=1}^{n} y_i^2 - b' \, \Sigma Xy}{n-1} = \frac{0.2292 - (0.49)(0.4668)}{4}$$

$$= 0.000111 \,,$$

$$s_{y|X} = 0.0105 \text{ volts.}$$

Certain assumptions are implied by the least-squares analysis as described:

(1) The function is linear. This seems reasonable from the plot in Figure J.1. However, the appropriateness of the linear function should be checked by techniques given in the literature.[1,2]

(2) The variation of the analyzer readings for a given level of concentration of NO_2 is the same for all levels. This is not usually a true assumption. Although there are insufficient data to check this assumption in this example, the assumption can be checked in practice by taking several observations at each concentration such as one would do in repeated zero-span checks. There are techniques[2] for fitting a weighted linear function and these can be employed when it is determined that the variations depend on the concentration levels.

(3) The levels of concentration of the standards are assumed to be precisely known. This is not exactly true as the standards are probably known to within 1 or 2 percent

(absolute). The important point here is that the concentrations should be known with much greater precision than the precision of the response. The effect of violations in the assumption on the above procedure is described in Reference 1 under the topic "Error in Both Variables."

In order to simplify the calculations and the discussion herein, it is assumed that the above assumptions apply and that no modified approach need be taken. These assumptions seem appropriate to illustrate the techniques to be described in this section and are adequate in most practical applications.

After the multipoint calibration curve has been determined, the curve is used to estimate the concentration of NO_2 from a given analyzer reading in volts for each sample until a recalibration is performed according to schedule or indicated need. The procedure to be used to make this decision will be discussed in Section J.3. The estimated precision of a reported concentration in ppm from a given analyzer reading involves an inverse prediction using the regression line of Figure J.2. That is, given an ordinate value y, what is the corresponding X? Note that the fitted line was derived on the assumption that given X, what is Y, the predicted mean value of y? The analyzer readings vary from replication-to-replication and day-to-day and thus there is an inherent error in these readings. This in turn is combined with the variation in the predicted values of y for given X as determined by the least-squares fitted line, to yield an estimate of the precision in the predicted concentration ob-

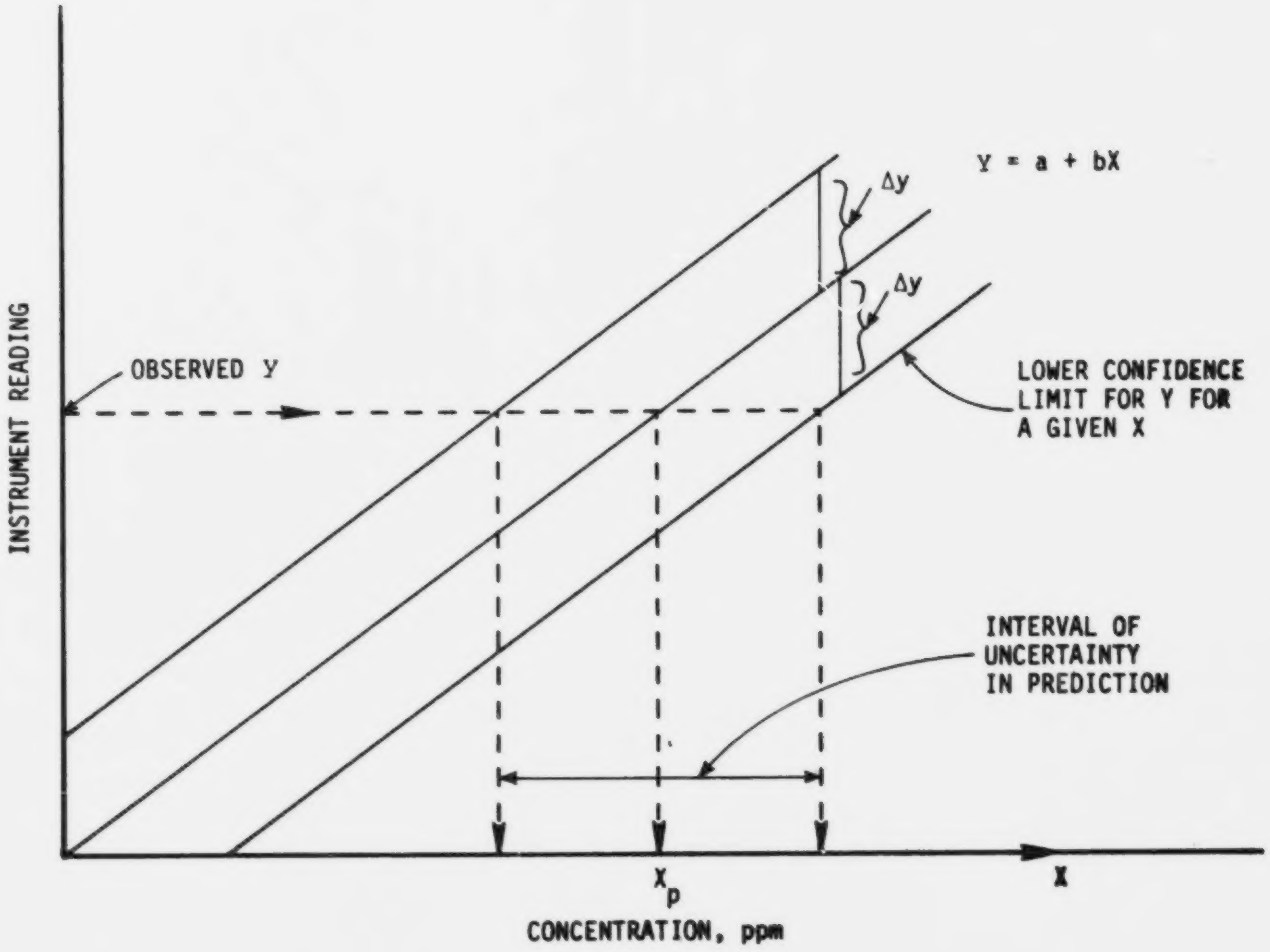

INSTRUMENT READING
OBSERVED Y
Y = a + bX
Δy
Δy
LOWER CONFIDENCE LIMIT FOR Y FOR A GIVEN X
INTERVAL OF UNCERTAINTY IN PREDICTION
X_p
X
CONCENTRATION, ppm

Figure J.2. Inverse prediction using regression line.

tained by inverting the equation $Y = -0.013 + 0.52X$ to obtain X_p the value of X predicted from the calibration curve,

$$X_p = \frac{Y + 0.013}{0.52}$$

or

$$X_p = 0.0258 + 1.938Y \text{ ppm}$$

thus for $y = 0.10$, $X_p = 0.22$ ppm. The question is, can an interval be given in the form that the true concentration falls between the limits (0.22 ± w) ppm? The answer is of course yes, but such an interval is rather tedious to calculate. Using the appropriate procedure[2], the interval for the concentration of NO_2 for $y = 0.10$ volts, is about 0.22 ± 0.07 ppm. Thus even though the calibration curve (line) appears to be reasonably precise, the resulting error in the predicted concentration can be quite large. This points out one of the important aspects in the calibration procedure, that is, relatively small errors in the deviations between the values observed and the calibrated values, can yield relatively large errors in predicted concentrations and thus this becomes one important source of uncertainty in the measurement process. The questions given in Section J.1, page 2 of 24 under MPC now become important in terms of how can these predictive errors be reduced to an acceptable order of magnitude, if this has not already been achieved. To reduce these errors, one or more of the following steps can be taken:

(1) Improve calibration standards (i.e., be sure that their levels are known as precisely as possible)

(2) Repeat the MPC, and use all of the data

(3) Adjust the spacing of the levels of the standards to yield improved precision within this expected range of the y's to be observed

(4) Repeat some levels of the MPC and use all of the data.

Consider here generally the effect of each of the above actions. First of all, action (1) will increase the cost of the acquisition and testing of the standards, and in maintaining and replacing of same. The tradeoff between these increased costs and the improvement which can be achieved would need to be considered by employing the techniques given under the subject "Error in Both Variables[1]." It is intuitively felt that if the stated concentrations of the standards are within 2 percent of their true values, then their overall impact on the precision of the predicted values will be sufficiently small.

Now consider what will happen if the MPC is repeated, action (2), (e.g. it may be redone on 5 consecutive days and the results put together to obtain one MPC based on 25 measurements, 5 at each level obtained on each of 5 days.) If the day-to-day variations are negligible, this procedure reduces the interval to about one-half the width for one MPC, i.e. the result could be reported as 0.22 ± 0.035 ppm instead of 0.22 ± 0.07 ppm as indicated for the previous example. In combining the results over 5 days, checks must be made of the consistency over the 5 days; if the analyzer behaved erratically on a specific day, this should be considered in the analysis, with the calibration data for that day possibly being discarded (See Appendix F). With additional repeated calibrations the interval width could be further reduced, but this approach is not suggested as a practical one for reasons to be given later.

The next consideration, action (3), is that of adjusting the spacing levels of the calibration gases. For example, if it is known (without any doubt) that the calibration curve is linear, then the best allocation of n levels of concentration would be as follows:

n even: make 50 percent of the observations at each end of the range of concentrations over which the linear relationship holds.

n odd: make 1 observation at the center as a confirmation of continued linearity and 50 percent of the remaining observations at each end of the range.

For additional information see a discussion of the optimal spacing of data for such problems.[3] If there is some doubt concerning the linearity assumption, the above approach would be poor and one should then take more center points or add two intermediate points to be able to check for nonlinearity. Some experience in the type of nonlinearity would be valuable in the selected spacing of the concentration of the calibration gases. This approach to improving the MPC and the resulting precision of the predicted concentrations is appropriate as considerable improvement in the precision can be obtained without repeating an entire MPC. This leads to the last and recommended procedure for improving almost all calibration procedures.

The fourth action suggests that certain levels of the MPC be repeated to improve on the precision of the results. It is beneficial to consider the combination of zero-span calibration or checks with the MPC in this approach. Hence before discussing this further, consider first the following discussion on OSC.

J.3 ZERO-SPAN CALIBRATION (OSC)

The zero-span calibration checks can be used to significantly improve on the MPC and to detect when a change may have occurred in the calibration, e.g. the instrument may have degraded because of a particular component wear-out, or the calibration gas may have degraded, changing concentration with time. These changes can be sudden (or catastrophic) or a slow degradation. The desire is to detect any changes that would significantly alter the calibration. The approach might be suggested that the daily OSC be used for that day and the equipment adjusted each day accordingly. However, this approach may result in predictive results with relatively poor precision as indicated in the previous discussion. Also, if the day-to-day variations are random variations in analysis which are consistent with the capabilities of the measurement process, then these calibration data should be combined to obtain an overall prediction with improved precision. Making daily adjustments when in fact they are not required only serves to increase the variation in the reported data. See Example J.2 in the following subsection. Before suggesting a calibration procedure, consider the questions posed in Section J.1, page 2 of 26, under OSC.

(1) How often do we make an OSC? This will be determined from the experience with day-to-day and within day variations in the OSC just as one needs to determine the causes of variability in a production process prior to determining the best sampling procedure to use in quality control checks. It would be suggested that initially they be made very frequently, e.g. at least once a day, until sufficient experience is gained, such as 25 to 30 OSCs should be required to test for trends or shifts in the levels.

(2) At which levels of the standards should the OSC be made? The answer to this question is similar to that point discussed under MPC, that is, if the calibration is truly linear, the precision can be increased by spreading out the levels of concentration. If however, the calibration is not linear the levels should be closer to the expected range of observed values of concentration for which the calibration curve will be used.

(3) What is the day-to-day and within day variation among the OSC checks? These variations are determined by using the same concentration levels on each day (or time of calibration check) and measuring the variation in the observed analyzer readings. These variations may depend on the level of the standard. For example, in the MPC example of Section J.2, the levels might be taken at 0.10 ppm and at 0.75 ppm. Calculate the value of s^2 for each set of measurements, i.e. analyzer readings at 0.10 ppm and at 0.75 ppm.

(4) How is a quality control chart used in determining if a significant drift or change in the calibration curve has occurred? The results obtained with each OSC can be plotted on a quality control chart for individuals, or results may be grouped for 2 or 3 days (OSC's) for checking on the range as well as the mean level. An example chart will be described in the following section.

(5) How does the expected range of concentration of the pollutant in the particular region (site being monitored) affect the selection of the OSC levels? The answer to this question would follow the recommendation under question (2) above.

J.4 SUGGESTED CALIBRATION PROCEDURE

It is not expected that a single proposed procedure could be appropriate for all calibrations used in ambient air pollutant measurement processes. However, it is felt that these procedures should follow the general set of guidelines set forth in this section. These guidelines provide some flexibility in the approach depending on the situation. Basically, it is assumed that an MPC can serve as a basis for predicting the concentration of a pollutant in an unknown sample over a period of several days or weeks and that the OSCs can be used to improve this prediction (i.e., its precision) until a significant change is detected in the measured values (by an NO_2 analyzer for example) in the zero-span gases. (The discussion herein can apply to calibration of a rotameter, pitot tube, etc.; however, the terminology usually used for gaseous calibration is used throughout this section to simplify the presentation and hopefully make it more readily understood). The significant change will be determined by an appropriate quality control chart or limits within which certain measurements or results should fall as previously suggested. Note that this approach accomplishes two goals simultaneously: (1) the OSC data can be combined with the MPC data to obtain an improved calibration curve so long as no significant change has been detected and,(2) if there is a significant change in the calibration, this will be detected in the same manner as in standard quality control procedures. If a point falls outside the prescribed control limits, the calibration should be checked and, if necessary, a new calibration gas may need to be acquired. By appropriate selection of the MPC levels, a check can be made initially on the linearity of the calibration. If it is considered to be linear, the OSC can be taken at two points spread out as far as reasonable in order to provide the desired precision. Thus after a few OSC's, in addition to a starting MPC, the combined data can quickly yield the desired precision in the estimated concentrations. An example was analyzed to illustrate the points made in the previous discussions. Only significant results are given herein as described under Example J.2.

Example J.2 - The data in this example were generated using the relationship that the analyzer reading y is a linear function of X plus some random error ε, i.e.,

$$y = \alpha + \beta X + 0.5u \ , \quad \varepsilon = 0.05u,$$

Example J.2

Calibration Form

Linear Regression (Least Squares)

(Straight line -- not through the origin)

Sample No.	X	y	Y(predicted mean)
1	______	______	
2	______	______	
3	______	______	
4	______	______	
5	______	______	
6	______	______	
7	______	______	
8	______	______	
9	______	______	
10	______	______	

ΣX = ______ Σy = ______

$\bar{X}$ = ______ $\bar{y}$ = ______

ΣX^2 = ______ Σy^2 = ______ ΣXy = ______

$(\Sigma X)^2/n$ = ______ $(\Sigma y)^2/n$ = ______ $(\Sigma X \cdot \Sigma y)/n$ = ______

Subtract

$\Sigma(X-\bar{X})^2$ = ______ $\Sigma(y-\bar{y})^2$ = ___ $\Sigma(X-\bar{X})(y-\bar{y})$ = ______

Divide by

$n-1$ = ______ $n-1$ = ______ $n-1$ = ______

s^2 = ______ s_y^2 = ______ s_{Xy} = ______

Slope of fitted line:

$b = s_{Xy}/s_X^2$ = ______

Equation of fitted line:

$Y = \bar{y} + b(X-\bar{X}) = (\bar{y}-b\bar{X}) + bX = a + bX$ = ______ + ______ X

Calculate predicted mean of Y for each X, record above. Variance of observed values (y) about fitted values (Y).

Σy^2 = ______ $(\Sigma y)^2/n$ = ______ $b\Sigma(X-\bar{X})(y-\bar{y})$ = ______

$s^2 = [\Sigma y^2 - (\Sigma y)^2/n - b\,\Sigma(X-\bar{X})(y-\bar{y})]/(n-2)$ = ______

$s = \sqrt{s^2}$ = ______ $s_b = \sqrt{\dfrac{s^2}{\Sigma(X-\bar{X})^2}}$ = ______

where u is a standard normal variable, i.e. with mean 0 and standard deviation 1. Consequently, y has a standard deviation of 0.05. For simplicity, α is assumed to be 0 and $\beta = 1$, thus

$$y = X + 0.05\mu$$

The values of μ were selected from a random normal variable table[4]. Five levels of X were assumed and resulting y values were determined in the initial multipoint calibration.

Multipoint Calibration Results

X	y
0.00	0.010
0.25	0.215
0.50	0.525
0.75	0.775
1.00	1.095

Zero-Span Results Each Day

X	y
0.10	0.075
0.75	0.805
0.10	0.035
0.75	0.73
0.10	0.155
0.75	0.76
0.10	0.115
0.75	0.765
0.10	0.155
0.75	0.765
0.10	0.145
0.75	0.780

The purposes of this example are to:

(1) Compare the predicted values of concentration based on observed analyzer readings where the predicted values are obtained by

a) Using only the initial MPC curve for each day,

b) Using the calibration curve based on all of the data given above,

c) Using a daily adjusted calibration curve based on data for the one day.

(2) Indicate the use of a QC chart for determining when a new MPC should be made.

(3) Compare predicted values where the calibration line is

a) Not required to pass through the origin,

b) Required to pass through the origin.

The comparison of the predicted values X_p with the known concentrations (based on assumed sampling procedure and equation of the line) is given in Table J.1 for cases (1) a, b, and c above. Comparing the upper part of the table with the lower part gives a comparison of the predictions under (3) a and b respectively. Even though only 6 days are used in the analysis, the results compare as would be expected, i.e. the best predictions are given by using all of the calibration data and the worst predictions result from daily adjustments. These results are summarized in the following table for easier comparison.

Table J.1 COMPARISON OF PREDICTIONS, $\hat{\sigma}$ *
BASIS OF PREDICTION

Line through origin	MPC	All data	Daily adjustment
No	0.033	0.030	0.038
Yes	0.040	0.030	0.042

These data are not sufficient to prove the value of different approaches, but are to illustrate the different procedures and fortuitously they also agree with theory in one respect, that is, using all the data is preferable. However a statistical comparison of these data would obviously not lead to any significant differences.

In conclusion, it is recommended that the following procedure be employed where feasible from a practical approach:

(1) MPC be done initially with several points, 4 is reasonable,
  a) To check for linearity,
  b) To determine an initial calibration curve,
  c) To estimate precision of data, analyzer values for given X values as measured by $s_{y|X}$

---

* $\hat{\sigma}$ is computed as the square root of the average sum of squares of deviations of the predicted values from the true values X_t, i.e. $\hat{\sigma} = [\frac{\Sigma(X_p - X_t)^2}{n}]^{1/2}$; (n=6 days in the example).

(2) Zero and span calibrations be done periodically, very frequently (or daily) at first and then less frequently depending upon drift, degradation, etc.

  a) Use 2 points (e.g. a zero and span) for linear calibration curve,

  b) Use 3 points otherwise, over a region sufficient to include expected concentration levels.

(3) QC chart be kept on the 2 points (the maximum and minimum points if 3 points are used) used in the zero and span calibrations,

  a) When no value falls outside the prescribed limits, use cumulative data to determine calibration, e.g. the calibration could be updated each week, month, etc. (A moving average curve could be used.)

  b) When a value falls outside the prescribed limits (UCL or LCL) rerun the MPC. If it does not differ significantly with the previous one, use all data to-date for an overall calibration. See Figure J.3. for a typical quality control chart.

  c) Because of the possible tendency of the zero and span values to change together or be correlated, a quality control chart may be maintained on the slope of the linear calibration (or the difference in the zero and span instrument readings, which is a constant multiple of the slope) and one point, say the span value. This quality control chart for the slope (or difference in

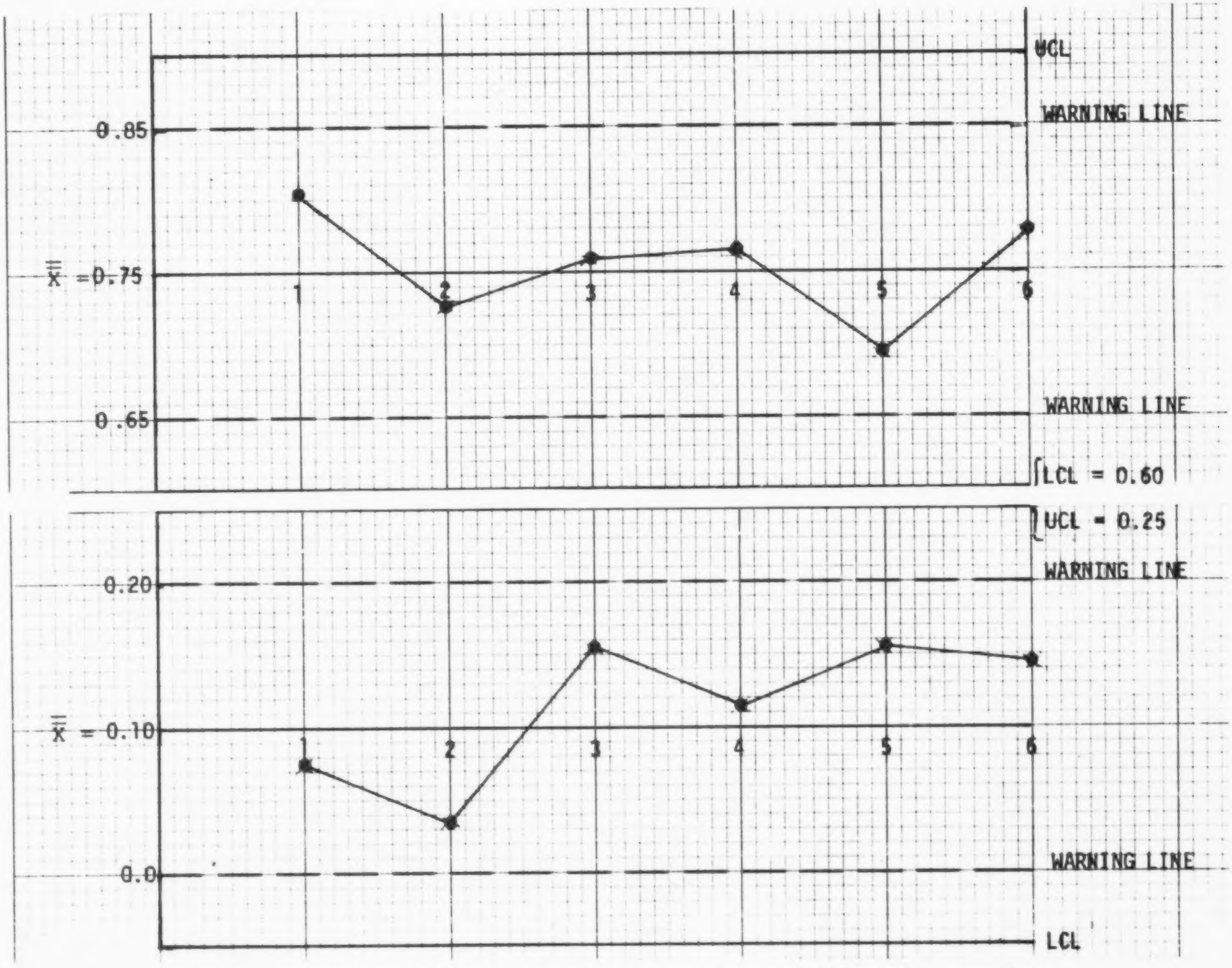


Figure J.3 Quality control charts for "zero" and span data.

zero and span values) may be more sensitive to changes in the instrument depending on the degree of correlation in the two readings.

J.5. REFERENCES

1. Mandel, John, The Statistical Analysis of Experimental Data, Interscience Publishers, John Wiley & Sons, Inc., New York, 1964.
2. Acton, F., Analysis of Straight Line Data, John Wiley & Sons, Inc., New York.
3. Ott, R. L., and Myers, R. H., Optional Experimental Designs for Estimating the Independent Variable in Regression, Technometrics, Vol. 10, No. 4, November, 1968.
4. Shewhart, W. A., Economic Control of Quality of Manufactured Product, Van Nostrand Co., New York, 1931.

APPENDIX K

INTERLABORATORY TESTS

K.1 CONCEPTS

One of the most commonly employed procedures to assess the quality of data being reported by a large number of agencies/laboratories...involved in a monitoring program is to conduct an interlaboratory test. Such a test typically involves preparing a set of samples, as nearly identical as possible, submitting them to a relatively large number of laboratories which are conducting the particular analysis, requesting that they be analyzed in the manner that a routine sample would be treated, and that the results be reported to the laboratory coordinating or conducting the test program.

One such interlaboratory procedure for performance of simulated ambient SO_2 bubbler samples is described in a recent report.[1] The reported results are analyzed to determine the variations among laboratories.

Another type of interlaboratory test program would be a collaborative test in which about ten to thirty labs might participate. The samples would be prepared in a similar manner to that in any interlab study, but in this case the analysis might be done on two or three days, duplicate or triplicate samples for each day. In this test, the data can be analyzed to determine not only the variation among laboratories, but also that within laboratories -- within days and among days.[2]

K.2 INTERLABORATORY TEST

The test program for simulated SO_2 bubbler samples is characteristic of almost any interlaboratory study; that is, samples were prepared at several levels of concentration (5 in this case) and sent to each laboratory for analysis. Thus the results might appear as follows where L laboratories participate in the test program.

	Concentration (Series)				
Laboratory	1000	2000	3000	4000	5000
1	42.68	112.66	241.04	310.40	374.00
2	51.70	135.00	290.00	350.00	478.00
3	34.72	107.00	190.00	301.30	248.00
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
L	56.50	133.33	276.37	351.60	427.06

Ideally each lab should report a value for each concentration level (or series).

In parallel with the test data provided by the individual labs, the laboratory coordinating the study should analyze (or have analyzed) each simulated sample with a sufficiently large number of analyses to estimate the mean and standard deviation of the concentration of the simulated samples with the desired precision. This provides an independent measure of the target value (overall mean) and variation in the simulated samples provided to the laboratories. These data can thus be used to determine limits within which the measurements obtained by the individual labs should fall with a given level of confidence.

K.2.1 Data Analysis - A brief discussion of the data analysis is given here in order that the participant/analyst in such programs can interpret the published results and the implication for his lab. There are several types of analyses given in the summary of results of such studies and this discussion will treat some of the more important aspects of these analyses, and refer the interested reader to some particular methods as described in the literature.[1,2,3] Clearly the sophistication of the analysis can vary considerably but the final results should be summarized so that they are useful to the laboratory analyst who wishes to determine whether his analysis procedure is satisfactory or not and also to put some limits on his own reported data. Thus the analysis should have three purposes:

(1) To summarize the overall test results.

(2) To provide the necessary information to each participant for use in describing the quality of his data. This information need be given only to each lab as it should not be important for each lab to know where another particular lab stands, only how all labs are doing as given by (1) above, in order that one lab may compare its results against all lab results.

(3) To provide each participant with a means of improving his analytical methods -- for example, preference for a certain method of analysis due to more accurate and precise results,

or identification of critical steps in the analysis by means of better written procedures. The analysis of the data only identifies the need for improving the method not usually the means by which this can be done.

The ultimate objective in interlab studies is to improve data quality as reported by many laboratories so that comparative studies can be made with the assurance that the data quality is sufficient to yield valid results and decisions.

K.2.1.1 Analysis of Interlaboratory Tests -

K.2.1.1.1 A summary analysis - As previously mentioned, one method of analysis of interlaboratory test data is described in an EPA report.[1] This provides an example of how the data from a large number of laboratories participating in an interlab test can be summarized. A summary analysis involves consideration of the following:

(1) Identification and elimination of outliers. There are two aspects to this analysis, determination whether a laboratory is an outlier, i.e. all of the observations from a laboratory indicate a significant bias (inaccuracy) and/or lack of precision in the measurement technique employed by a lab, and whether a single measurement by a lab is an outlier. Various test procedures are applied in the literature for identifying individual outliers, e.g. see the references at the end of Appendix F. Similarly, there are alternative pro-

cedures for checking the consistency of one lab versus all other labs. Some of these procedures are described in the referenced report.[1] Some other procedures are given in appropriate statistical texts.[3,4]

(2) Estimation of within laboratory and among laboratory variation. After the elimination of outliers (laboratories or individual measurements) by some appropriate method, it is then desired to summarize the remaining data by means of a measure of variation of the results within a laboratory and the variation among laboratories. The latter variation can be considered a measure of the variation of the inaccuracies or biases among the participating labs. Because the variation is typically dependent on the concentration level, it is usually desired to conduct the analysis in a manner to make the appropriate estimates. For example, the analysis can be separately performed for each concentration level, or by making a transformation of the data.[5]

K.2.1.1.2 Analysis for a particulate lab - A second purpose for the analysis is to provide the results in a manner as to be useful to a participant in the test program, that is, so that the lab may evaluate if its analysis procedure is consistent with the results of all laboratories excluding outliers and assess the quality of its reported data. Suppose that the results of two particular labs are as indicated in Table K.1 and that the results of all labs are summarized by their mean $\overline{X}$ and standard deviation s. Also assume for convenience that n=50 values are used to calculate $\overline{X}$ and s.

Table K.1 COMPARISON OF INDIVIDUAL LAB RESULTS WITH ALL LAB RESULTS

Data	Series				
	1000	2000	3000	4000	5000
Lab X	43.2	118.3	255.3	351.6	374.4
$(X-\overline{X})/s$	- 0.03	0.05	0.17	0.67	- 0.29
Lab Y	58.3	132.0	270.0	361.0	560.1
$(Y-\overline{X})/s$	1.0	0.88	0.59	0.89	3.56
All Labs (n)	50	50	50	50	50
$\overline{X}$	43.7	117.4	249.5	322.8	388.2
s	14.6	16.6	34.8	43.0	48.3
$\hat{CV} = 100\ \overline{X}/s$	33.4	14.2	13.9	13.3	12.4

If n is reasonably large, then $\frac{X - \overline{X}}{s}$* is approximately normally distributed with mean zero and standard deviation one. Hence this value can be plotted on a quality control chart with limits at UCL = 3, LCL = -3, $\overline{\overline{X}} = 0$. The value for each concentration, five in all, would be plotted at each sample number, or one can plot their range on a range chart and their mean on a mean chart with limits given by:

---

*Actually, $(X-\overline{X})/s\sqrt{1 - \frac{1}{n}}$ has a t distribution where X is one of the measurements of the sample from which the sample mean $\overline{X}$ is determined.

Range chart: $\bar{R} = 2.3$, $UCL_R = D_4\bar{R} = 4.91$, $LCL_R = 0$

Mean Chart: $\bar{\bar{X}} = 0$, $UCL_{\bar{X}} = \frac{3}{\sqrt{5}} = 1.34$,

$$LCL_{\bar{X}} = - \frac{3}{\sqrt{5}} = 1.34$$

The plot of the individuals can reveal some useful information; particularly, if the values are identified by concentration level. For example, a relationship between $\frac{X-\bar{X}}{s}$ and concentration may indicate a poor calibration (error in slope).

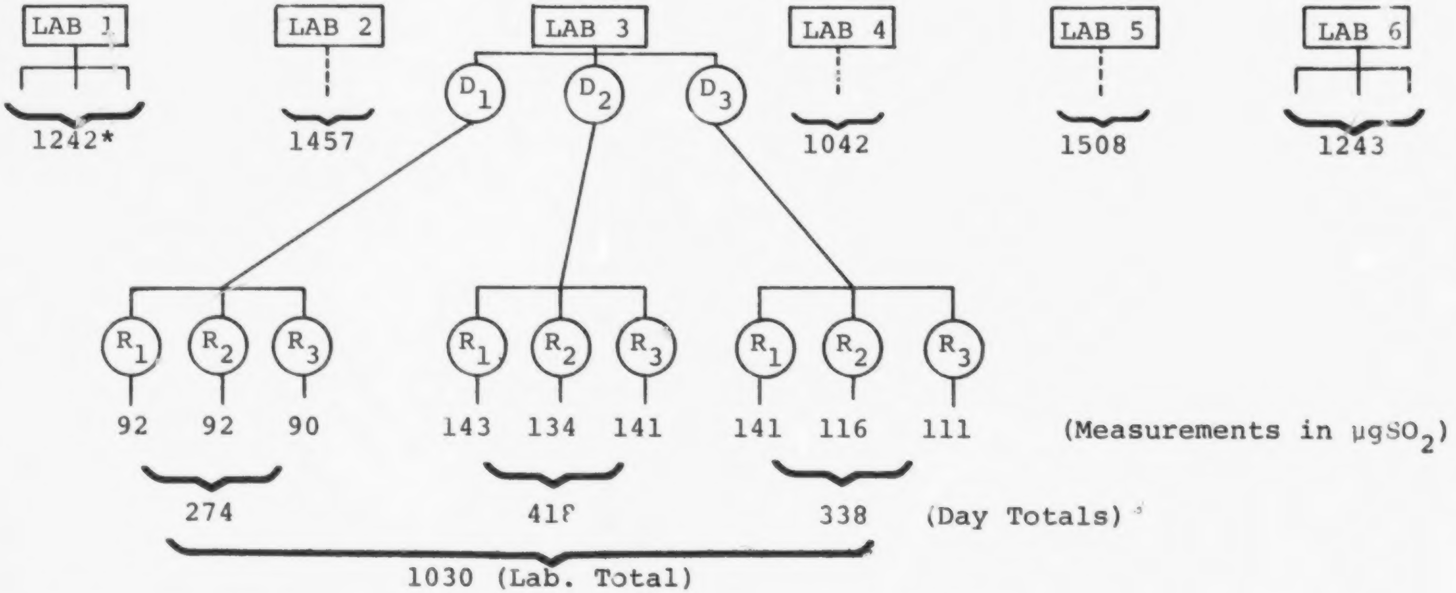
For the hypothetical laboratories given in Table K.1, one would infer that Lab X has obtained values consistent with the overall lab average, whereas Lab Y has tended to be high and for Series 5000 was significantly higher or "out of line" with respect to the mean for all labs.

K.2.1.1.3 Feedback of information on methods - In order to attain the overall purpose of improvement in data quality, there needs to be a feedback of information to the laboratories having some difficulties with the measurement process. This can be done, for example, through better method description, ruggedness tests to determine the critical steps in the method (this really should have been done prior to extensive use of a method), and/or consultation with experienced and well-qualified analysts.

K.2.1.2 Analysis of Collaborative Test Data - The analysis of collaborative test data involves a more detailed breakdown of the total variation in measurements, i.e. not only the variation among labs (reproducibility), but an estimate of the variation among days within labs (repeatability) and among replicates within days (replicability). A partial selection of data from a collaborative test study as given in Figure K.1 indicates the pattern of variation which is reasonably typical of such test data. For such data, the smallest component of variation is expected for within day or a measure of replicability, the next largest for among days or repeatability, the largest being among labs or reproducibility. These three components of variation are denoted by σ_r^2, σ_d^2, and σ_ℓ^2 which are the variances for the replicates, days, and labs respectively. These three components are then combined to obtain the three measures of variation.

Measure	Variance	Standard Deviation
Replicability	σ_r^2	σ_r
Repeatability	$\sigma_r^2 + \sigma_d^2$	$\{\sigma_r^2 + \sigma_d^2\}^{1/2}$
Reproducibility	$\sigma_r^2 + \sigma_d^2 + \sigma_\ell^2$	$\{\sigma_r^2 + \sigma_d^2 + \sigma_\ell^2\}^{1/2}$

Thus a measure of replicability is the variance or standard deviation among measurements made in the same day by the same analyst in the same laboratory. Repeatability is the sum of


LAB 1
LAB 2
LAB 3
LAB 4
LAB 5
LAB 6
D_1
D_2
D_3
1242*
1457
1042
1508
1243
R_1
R_2
R_3
R_1
R_2
R_3
R_1
R_2
R_3
92
92
90
143
134
141
141
116
111
(Measurements in $\mu g SO_2$)
274
418
338
(Day Totals)
1030 (Lab. Total)

* Only the lab total is given for labs other than LAB No. 3.

Figure K-1 Data selected from a collaborative test study[2] ($\mu g SO_2$)

the replicate (within day) and the day-to-day variance for the same laboratory. Finally, reproducibility is the measure of total variation across days and laboratories. Typically the square roots of these variances, i.e. the standard deviations, are used as the corresponding measures. In some references, the measures are defined as the upper confidence limit for the difference between two repeat measurements which would be exceeded at most 10 percent (1 percent) of the time.

The specific method of analysis is described in most statistical texts on design and analysis of experiments under the subject of nested experiments and components of variance.

K.3 REFERENCES

1. Bromberg, S.M., et al, Survey of Laboratory Performance Analysis of Simulated Ambient SO_2 Bubbler Samples, Survey No. 2, Sept. 1964, U.S. Environmental Protection Agency.

2. McKee, H.C., Childers, R.E., Saenz, O.S., Jr. Collaborative Study of Reference Method for the Determination of Suspended Particulates in the Atmosphere (High Volume Method), Contract CPA 70-40, SwRI Project 21-2811, June 1971.

3. Mandel, J. The Statististical Analysis of Data, Interscience Publishers, a division of John Wiley and Sons, Inc., New York, 1964.

4. Youden, W.J., Statistical Techniques for Collaborative Tests, The Association of Official Analytical Chemists, Box 540, Benjamin Franklin Station, Washington, DC 20044. (Price $3.00)

5. Explanation of the NIOSH PAT Statistical Protocol, NIOSH, Cincinnati, OH. (PAT-- proficiency analytical testing is the NIOSH interlaboratory program for assessing the quality of analytical data.)

APPENDIX L

RELIABILITY AND MAINTAINABILITY

L.1 INTRODUCTION

Reliability and maintainability are those aspects of quality assurance which are concerned with the quality of performance of a component, instrument, or a process over time. One definition of reliability is--the probability of performing a specified function without failure under given conditions for a specified period of time or cycles of usage. The common statistical meaning of the term reliability is that quality that a process has of producing consistent and dependable results.

Maintainability is defined as--the probability that given a component, instrument, or process has failed, the component will be repaired or replaced and thus become operational within a specified time. An easily maintained system would be one for which troubleshooting for failures and the necessary repair or replacement of components can be performed in a short time. Diagnostic troubleshooting procedures provided by the manufacturer need to be included in the operator's manual.

Maintenance for automatic monitoring equipment can be divided into two types: routine (preventive) and nonroutine (corrective). Routine maintenance procedures are necessary to provide optimum operational performance and minimum instrument downtime. Nonroutine maintenance (corrective) is performed to rectify instrument failures or degraded performance of the instrument or to repair any part of the total system. Preventive maintenance should decrease the need for corrective maintenance.

An overall measure of system performance over time is availability. This can be simply defined as the ratio of the uptime of the system to the sum of uptime and downtime, i.e. Thus A is a measure of the ability of the system to perform its intended function when called upon to do so,

$$A = \frac{U}{U+D}$$

where A = availability of the system

U = uptime

D = downtime

Note that U+D is not necessarily the total calendar time, this would only be so if the system were used 24 hours per day every day. If the system is being used every third day, preventive maintenance can be performed on the off-day, not decreasing the system availability. Thus the availability is increased by having a reliable system (few failures) or a highly maintainable system (ease of maintenance) or a combination of these. In the pollutant monitoring systems it is desired that a system provide consistent and dependable data over long periods of time or continuously, thus a high availability is necessary. See Figure L.1 for a flow diagram of reliability and maintainability actions.

The causes of unreliability of an instrument are many. The complexity of measuring instruments made up of many components may result in failure if any one component fails. Hence the likelihood of failure increases rapidly with the increasing number of components which need to perform a specified function in order that the system performs its required function. Records need to be kept on each component so that the time-to-failure (or degraded performance), cause of failure, and maintenance time can be determined for future use in procurement and maintenance practices. Clearly unreliable equipment or poorly maintainable equipment are to be eliminated from consideration in future procurement.

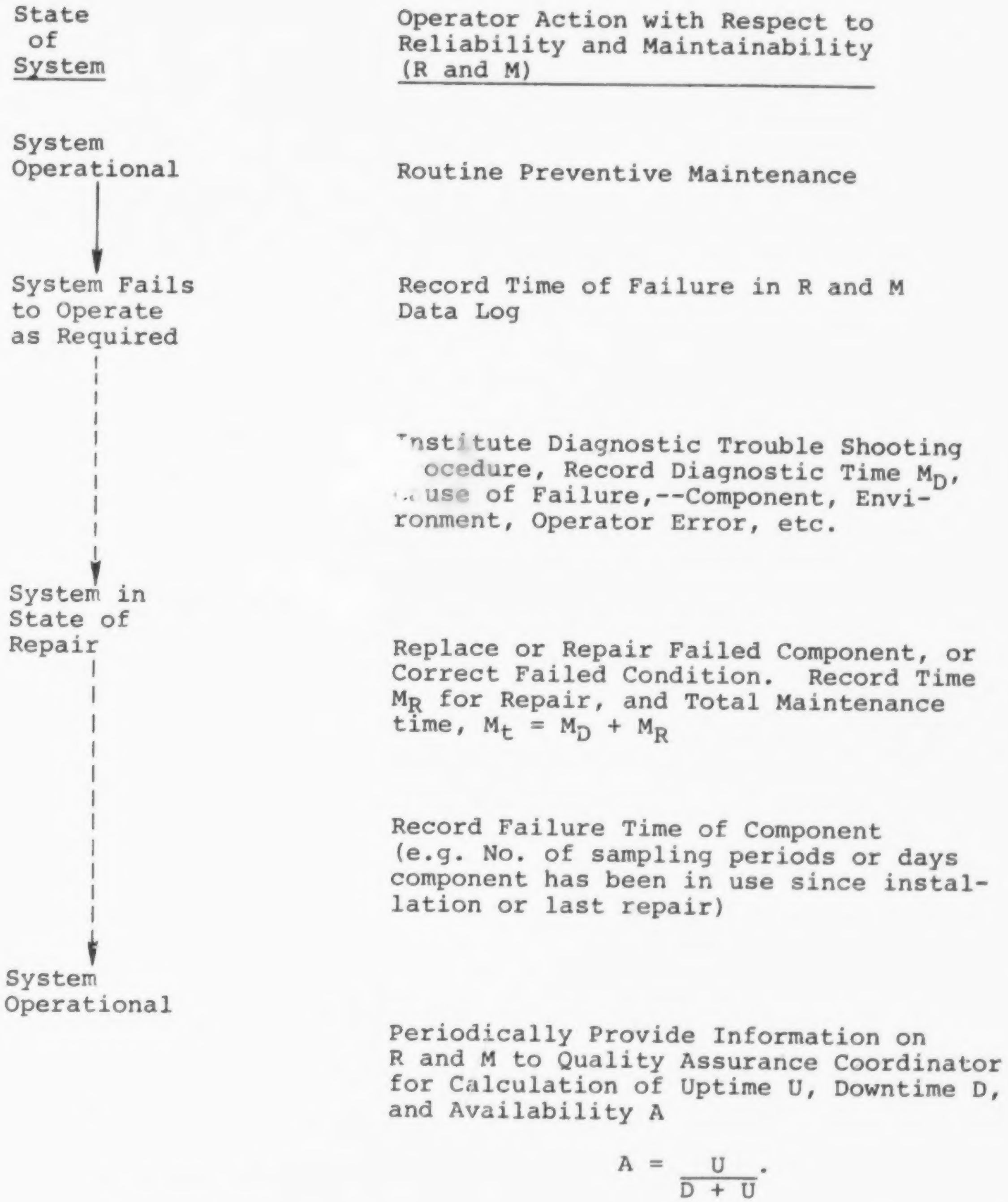
Field operational reliability depends upon three important variables in addition to the reliability of the components. Thus the reliability of the entire monitoring system is determined by the proper combination of these separate reliabilities:

(1) Environment or operating conditions.

(2) Functional interactions between various subsystems.

(3) Conditions imposed by the operators.

An analytical calculation of reliability requires a great deal of knowledge concerning these factors. However, an empirical estimate can be obtained if the proper failure data are recorded. It is important to realize that some systems may be unreliable in some environments,



State of System	Operator Action with Respect to Reliability and Maintainability (R and M)
System Operational	Routine Preventive Maintenance
System Fails to Operate as Required	Record Time of Failure in R and M Data Log Institute Diagnostic Trouble Shooting Procedure, Record Diagnostic Time M_D, Cause of Failure,--Component, Environment, Operator Error, etc.
System in State of Repair	Replace or Repair Failed Component, or Correct Failed Condition. Record Time M_R for Repair, and Total Maintenance time, $M_t = M_D + M_R$ Record Failure Time of Component (e.g. No. of sampling periods or days component has been in use since installation or last repair)
System Operational	Periodically Provide Information on R and M to Quality Assurance Coordinator for Calculation of Uptime U, Downtime D, and Availability A

$$A = \frac{U}{D + U}.$$

Figure L.1 Flow diagram of reliability and maintainability actions.

e.g. high humidity/high ambient temperature, whereas other systems may be unreliable if the operator is not properly trained.

Figure L.2 illustrates the typical life history of some types of equipments. The first phase is a debugging phase, the presence of marginal or short-life parts at original operation (burn-in) is characterized by a decreasing hazard rate. The higher hazard rate period is commonly called the infant mortality period, and it results primarily from poor or marginal workmanship. The next phase is characterized by a relatively constant hazard rate which is the effective life of the equipment. This is followed by a period of increasing hazard rate, the beginning of wear-out failures in the equipment. This curve demonstrates the need for

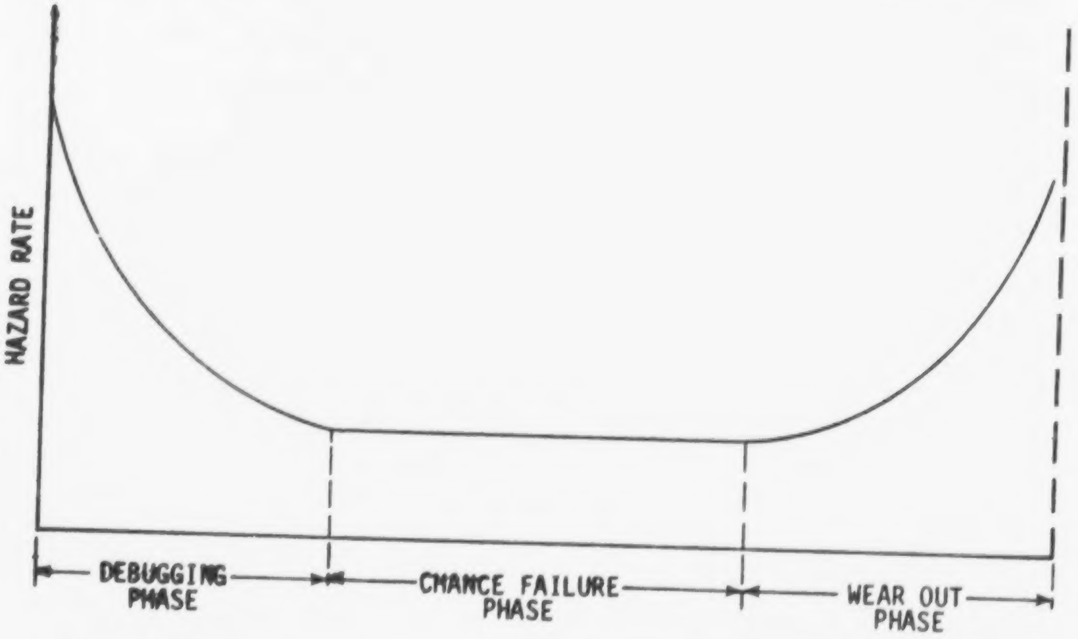


Figure L.2 Life history of some types of equipment.
(the bathtub curve)

preventive maintenance in equipment exhibiting a history of behavior as illustrated in Figure L.3. If the debugging phase is short and if the hazard rate is relatively high, the equipment should be (1) tested through the burn-in period, or (2) the short-lived components should be identifed and the appropriate equipment design modifications implemented. There are several techniques which can be employed to improve the reliability/availability of equipment. Some of these are listed here with no discussion. Reference to appropriate books on reliability would be necessary to select and apply the best technique.[1,2,3]

(1) Design should be simple as possible.

(2) Derating is used to achieve higher reliability (safety margin).

(3) Burn-in if decreasing hazard rate in early life.

(4) Redundancy (active, standby, and spares).

(5) Protection from environmental or operational stresses--packaging and handling.

(6) Ease of maintenance, service (actually an aid to increasing availability not reliability).

(7) Use of appropriate screening tests for parts, components.

(8) Specify replacement schedules (increases reliability by preventive maintenance).

(9) Conduct reliability qualification/demonstration tests.

Designing for reliability requires a cost-trade-off between unit cost of failure and the unit investment in reliability as illustrated by Figure L.3.

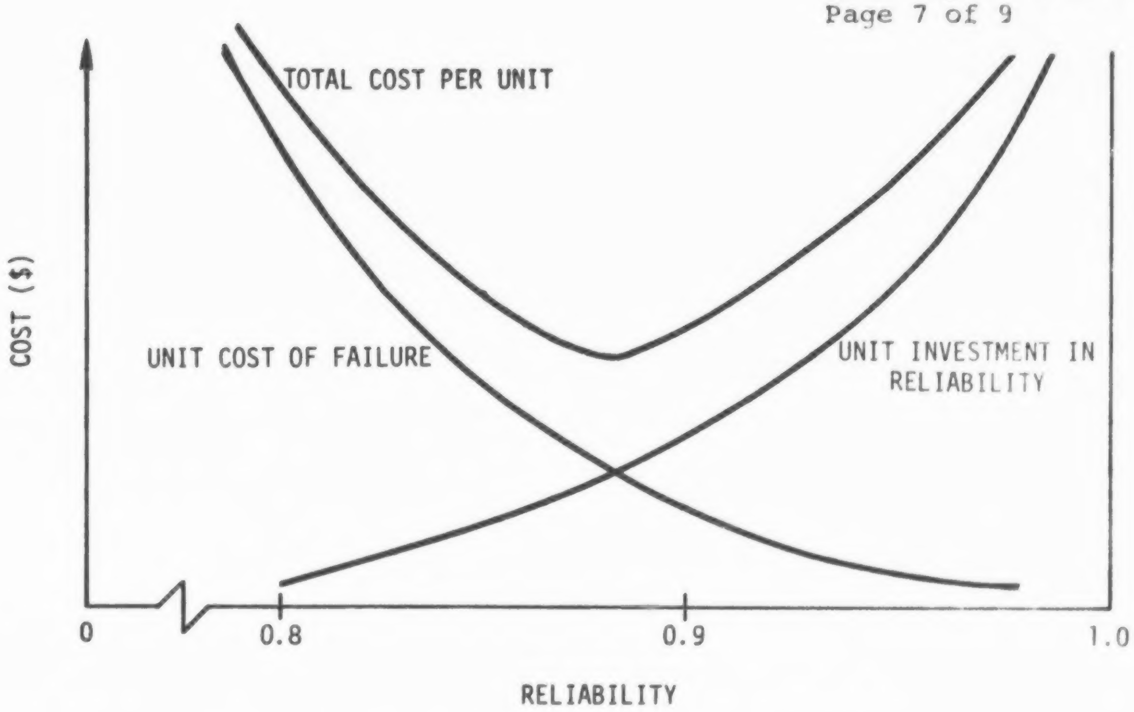


Figure L.3 Reliability cost-trade-off considerations.

L.2 LIFE TESTING

The life of an item is an important characteristic to the user and consequently life testing is often specified to demonstrate that a part component or system has a mean life of a required number of hours. In order to assure with high confidence that a system or item is acceptable it may be required that n systems be tested for example, until either

(1) R systems have failed

(2) The test time does not exceed a specified time, e.g. 1000 hours,

i.e the test may be failure terminated (1), or time terminated (2), above. The number of failures r may be put equal to n, in which case the test is run until all items have failed. This is usually not practical due to the expense of testing and the consequent time delay in reaching a decision.

The definition of failure may be a sudden discontinuity of operation or a gradual degradation of the performance of the item as measured by pertinent performance parameters. Life of an item is thus defined as the time that the item continues to perform satisfactorily under specified conditions of operation.

The information required to plan a life test consists of

(1) The sample size - number of items to be placed on test.

(2) The testing approach (whether failure-or time-terminated.

(3) The decision of whether to replace failed items.

(4) The termination criteria (failure criteria).

(5) The degree of precision (e.g. the desired statistical confidence in the case of estimating the mean life of the item).

(6) The model assumed for the distribution of failure times, exponential, Weibull, lognormal, etc.

Test plans have been developed for life testing (acceptance sampling) when the failure time distribution is the negative exponential (MIL-STD-781) and other distributions. The negative exponential distribution implies a constant hazard rate, i.e. the time-to-failure of an item is independent of the time that it has been on test. On the other hand, the Weibull distribution can be applied when the item has a decreasing hazard rate (early life, debugging phase) or increasing hazard rate (wearout) phase at end of life), see the bathtub curve of Figure L.1. It is not possible to treat the many problems of testing in this short appendix, thus the reader is referred to one of the many texts on the subject.[1.4]

L.3 REFERENCES

1. Juran, J.M., Quality Control Handbook, Third Edition, McGraw Hill Book Co., New York, Section 8. (Also see Section 25, pp 25-41, for a discussion of life testing and reliability.

2. Enrick, N.L., Quality Control and Reliability, Sixth Edition, Industrial Press, Inc., New York, 1972.

3. Myers, R. H., Wong, K. L., and Gordy, H.M., Reliability Engineering for Electronic Systems, John Wiley & Sons, Inc., New York, 1964.

4. MIL-STD-781, Reliability Tests, Exponential Distribution, 1967, Department of Defense, Washington, D.C.